



(19) **BUNDESREPUBLIK  
DEUTSCHLAND**



**DEUTSCHES  
PATENT- UND  
MARKENAMT**

(12) **Offenlegungsschrift**  
(10) **DE 197 54 622 A 1**

(21) Aktenzeichen: 197 54 622.6  
(22) Anmeldetag: 9. 12. 97  
(43) Offenlegungstag: 10. 6. 99

(51) Int. Cl.<sup>6</sup>:  
**C 12 N 9/10**  
C 12 N 15/54  
C 07 H 21/04  
C 12 N 15/63  
C 07 K 16/40  
C 12 N 1/00  
C 12 N 15/82  
A 01 H 5/00  
A 01 H 1/00  
C 12 N 5/10  
C 07 K 14/415

**DE 197 54 622 A 1**

(71) Anmelder:  
Schawen, Antje von, Dr., 49076 Osnabrück, DE  
  
(74) Vertreter:  
WUESTHOFF & WUESTHOFF Patent- und  
Rechtsanwälte, 81541 München

(72) Erfinder:  
gleich Anmelder  
  
(56) Entgegenhaltungen:  
WO 96 21 038  
CA 119:245692f;  
CA 120:294245s;  
EMBL-Genbank AC B24856;  
EMBL-Genbank AC AC000098;

**Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen**

Prüfungsantrag gem. § 44 PatG ist gestellt

- (54) Pflanzliche Gntl-Sequenzen und Verwendung derselben zur Gewinnung von Pflanzen mit verminderter oder fehlender N-Acetylglucosaminyltransferase I (GnTI)-Aktivität
- (57) Die Erfindung betrifft pflanzliche Gntl-Sequenzen, insbesondere pflanzliche Nukleinsäuresequenzen, die das Enzym N-Acetylglucosaminyltransferase I (GnTI) kodieren, davon abgeleitete DNA-Sequenzen einschließlich Gntl-"antisense"- und "sense"-Konstrukten, und deren Translationsprodukte, gegen diese Translationsprodukte gerichtete Antikörper sowie die Verwendung der Sequenzinformation zur Gewinnung von transformierten Mikroorganismen und von transgenen Pflanzen, einschließlich solchen mit verminderter oder fehlender N-Acetylglucosaminyltransferase I-Aktivität. Derartige Pflanzen mit verminderter oder fehlender N-Acetylglucosaminyltransferase I-Aktivität sind zur Herstellung von Glykoproteinen spezifischer Konstitution bezüglich der Zuckerreste von großer Bedeutung.

**DE 197 54 622 A 1**

Die Erfindung betrifft pflanzliche GntI-Sequenzen, insbesondere pflanzliche Nukleinsäuresequenzen, die das Enzym N-Acetylglucosaminyltransferase I (GnTI) kodieren, wie auch davon abgeleitete GntI-"antisense"- bzw. "sense"-Konstrukte, und deren Translationsprodukte, gegen diese Translationsprodukte gerichtete Antikörper sowie die Verwendung der Sequenzinformation zur Gewinnung von transformierten Mikroorganismen und von transgenen Pflanzen, einschließlich solchen mit verminderter oder fehlender N-Acetylglucosaminyltransferase I-Aktivität. Derartige Pflanzen mit verminderter oder fehlender N-Acetylglucosaminyltransferase I-Aktivität sind zur Herstellung von Glykoproteinen spezifischer Konstitution bezüglich der Zuckerreste von großer Bedeutung.

#### Stand der Technik

In Eukaryonten werden Glykoproteine im Endoplasmatischen Retikulum (ER) cotranslational (d. h. bei Import in das ER-Lumen) durch Verknüpfung von zunächst membrangebundenen Glykanen (an Dolicholpyrophosphat) mit spezifischen Asparagin-Resten in der wachsenden Polypeptidkette zusammengesetzt (N-Glykosylierung). In höheren Organismen unterliegen Zuckereinheiten, die an der Oberfläche der gefalteten Polypeptidkette liegen, in den Golgi-Zisternen weiteren Trimm- und Modifikationsreaktionen (Ref. 1). Durch verschiedene Glykosidasen und Glykosyltransferasen im ER werden zunächst typische GlcMan<sub>5</sub>GlcNAc<sub>2</sub>-Grundeinheiten des "high"-Mannose-Typs gebildet und anschließend bei Passage durch die verschiedenen Golgi-Zisternen in sogenannte "komplexe" Glykane umgewandelt. Letztere zeichnen sich durch weniger Mannose-Einheiten und den Besitz weiterer Zuckerreste, wie Fucose, Galaktose und/oder Xylose in Pflanzen bzw. Sialinsäure (N-Acetylneuraminsäure, NeuNAc) in Säugern, aus (Ref. 1, 2, 3). Das Ausmaß der Modifikationen ist von Glykoprotein zu Glykoprotein verschieden. Einzelne Polypeptidketten können heterogene Zuckerketten tragen. Zudem kann das Glykosylierungsmuster für ein bestimmtes Polypeptid variieren (gewebespezifische Unterschiede), und muß auch nicht immer bezüglich einer bestimmten Glykosylierungsstelle gleich sein, was als "Mikroheterogenität" bezeichnet wird (Ref. 4, 5). Die Rolle von Asparagin-ständigen Glykanen ist bislang kaum verstanden, was u. a. daraus resultiert, daß diese mehrere Funktionen erfüllen können (Ref. 6). Jedoch ist anzunehmen, daß beispielsweise der Schutz einer Polypeptidkette vor proteolytischem Abbau auch durch Glykane eines einfacheren Oligomannosyl-Typs gewährleistet werden kann (Ref. 7).

#### Problemstellung

Glykoproteine sind für Medizin und Forschung von großer Bedeutung. Eine Isolierung von Glykoproteinen im Großmaßstab ist jedoch aufwendig und teuer. Die direkte Anwendung konventionell isolierter Glykoproteine ist oft problematisch, da einzelne Reste der Glykan-Komponenten bei Verabreichung als Therapeutikum ungewünschte Nebenwirkungen auslösen können. Dabei trägt die Glykan-Komponente vor allem zu den physikochemischen Eigenschaften (wie Faltung, Stabilität und Löslichkeit) der Glykoproteine bei. Des weiteren tragen isolierte Glykoproteine, wie bereits ausgeführt, selten einheitliche Zuckerreste, was als "Mikroheterogenität" bezeichnet wird.

Hefen erweisen sich zur Gewinnung von Glykoproteinen für Medizin und Forschung als ungeeignet, da sie nur Glykosylierungen zum sogenannten "high"-Mannose-Typ durchführen können. Insekten und höhere Pflanzen zeigen zwar "komplexe", aber von Tieren abweichende Glykoproteinmodifikationen. Aus Pflanzen isolierte Glykoproteine wirken deshalb in Säugern stark antigen. Tierische Organismen mit Glykosylierungsdefekten sind meist nicht lebensfähig, da die terminalen Glykan-Reste (beispielsweise membranständige Glykoproteine) meist biologische Signalfunktion besitzen und vor allem für die Zell-Zellerkennung während der Embryonalentwicklung unentbehrlich sind. Es existieren zwar bereits Säugerzelllinien mit definierten Glykosylierungsdefekten, jedoch ist deren Kultivierung arbeitsintensiv und teuer.

Im Einzelnen wurden für Säuger auf Zellkulturebene bereits unterschiedliche Glykosylierungsmutanten beschrieben (Ref. 7, 8, 9, 10). Diese Mutanten sind entweder in der Biosynthese reifer Oligosaccharidketten am Dolicholpyrophosphat oder in der Glykan-Prozessierung betroffen bzw. zeigen Abweichungen in ihren terminalen Zuckerresten. Einige dieser Zelllinien weisen einen konditional-lethalen Phänotyp auf oder zeigen Defekte im intrazellulären Proteintransport. Die Folgen dieser Defekte für den intakten Organismus sind schwer abschätzbar. Es wurde beobachtet, daß eine Veränderung im Muster komplexer Glykane auf Zelloberflächen von Säugern mit Tumor- und Metastasenbildung einhergeht, obwohl eine funktionale Beziehung noch nicht eindeutig nachgewiesen werden konnte (Ref. 9). Glykosylierungsmutanten kommen in Säugern daher sehr selten vor. Diese unter HEMPAS (Hereditary Erythroblastic Multinuclearity with a Positive Acidified Serum lysis test) zusammengefaßten Defekte (Ref. 10, 11) beruhen entweder auf einem Defizit von Mannosidase II und/oder niedrigen Gehalten des Enzyms N-Acetylglucosaminyltransferase II (GnTII) und wirken sich stark einschränkend auf die Lebensfähigkeit des mutierten Organismus aus. GntI-"knock out"-Mäuse, in denen das Gen für GnTI zerstört wurde, sterben bereits "in utero" an multiplen Entwicklungsdefekten (mündliche Mitteilung, H. Schachter, Toronto).

Für Pflanzen waren bis vor kurzem keine vergleichbaren Mutanten bekannt. Durch den Einsatz eines Antiserums, das spezifisch "komplex"-modifizierte Glykanketten pflanzlicher Glykoproteine erkennt und hauptsächlich gegen die hochantigenen  $\beta$ 1 $\rightarrow$ 2 verknüpften Xylose-Reste gerichtet ist (Ref. 12), konnte die Anmelderin aus einer EMS-mutagenisierten F<sub>2</sub>-Population der genetischen Modellpflanze *Arabidopsis thaliana* mehrere unabhängige Mutanten isolieren, die keine "komplexe" Glykoproteinmodifikation mehr zeigten (complex glycan, cgl-Mutanten). Nach mindestens fünf Rückkreuzungen, jeweils gefolgt von intermittierenden Selbstungen (zum Wiederauffinden des rezessiven Defekts), wurden die Glykoproteine analysiert. Diese trugen hauptsächlich Glykane des Man<sub>5</sub>GlcNAc<sub>2</sub>-Typs, was auf einen Defekt in N-Acetylglucosaminyltransferase I (GnTI) hindeutete (Ref. 8). Tatsächlich fehlte den *Arabidopsis* cgl-Mutanten GnTI-Aktivität (Ref. 13), welche normalerweise die erste Reaktion im Syntheseweg zu "komplex" modifizierten Zuckerketten katalysiert (Ref. 1). Nach bisherigen Beobachtungen resultiert daraus allerdings keine Beeinträchtigung der Lebensfähigkeit der mutierten Pflanzen. In neueren Publikationen werden Pflanzen als mögliche Quelle zur Herstellung

von pharmazeutisch relevanten Glykoproteinen oder Vakzinen vorgeschlagen (Ref. 14, 15). Darin wird jedoch übersehen, daß aus Pflanzen isolierte Glykoproteine in Säugern starke Immunreaktionen auslösen können, was bisher einer Produktion heterologer Glykoproteine in Kulturpflanzen im Wege stand.

Pflanzen kommen weitgehend ohne "komplex" modifizierte Glykoproteine aus, wie die Anmelderin am Beispiel der Arabidopsis egl-Mutante zeigen konnte (Ref. 13). In der Mutante werden sekretorische Proteine im ER zunächst normal glykosyliert. Im Golgi-Apparat der Cgl-Mutante bleiben die über Asparagin-Reste (N-Glykosylierung) an das Polypeptidrückgrat gebundenen Oligomannosylketten dann jedoch auf Stufe von Man<sub>5</sub>GlcNAc<sub>3</sub>-Resten stehen, da N-Acetylglucosaminyltransferase I (GnTI) -Aktivität fehlt (Fig. 1). Durch diesen Biosyntheschloß wird die pflanzenspezifische "komplexe" Glykoproteinmodifikation und insbesondere die Anheftung von  $\alpha$ 1 $\rightarrow$ 3 Fucose- und  $\beta$ 1 $\rightarrow$ 2 Xylose-Resten verhindert, wodurch die stark antigene Wirkung auf den Säugerorganismus entfällt. Als krautige Pflanze besitzt Arabidopsis jedoch wenig verwertbare Biomasse. Zur Herstellung von biotechnologisch relevanten Glykoproteinen im Großmaßstab sind diese egl-Pflanzen deshalb weniger geeignet. Als Alternative wären Kultursorten, insbesondere Solanaceen, wie z. B. Kartoffel, Tabak, Tomate oder Paprika, und des weiteren Luzerne, Raps, Rüben, Soja, Salat, Mais, Reis, und Getreide, mit fehlender oder stark gedrosselter GnTI-Aktivität ideal zur Produktion von heterologen Glykoproteinen in Pflanzen. Dazu würden sich Verfahren des "homology-dependent gene silencing" (Ref. 16, 17) anbieten.

Wie Fig. 3 zeigt, ist die Homologie der ersten ermittelten pflanzlichen GnTI-Sequenz aus Kartoffel (*Solanum tuberosum* L., St) im Vergleich mit den entsprechenden bekannten Sequenzen aus tierischen Organismen außerordentlich niedrig (nur 30-40% Identität auf Proteinebene, vgl. Fig. 3A), so daß eine effiziente Drosselung der endogenen "komplexen" Glykoproteinmodifikation in Pflanzen mittels "antisense" bzw. "sense"-Suppression (Ref. 21) durch die Verwendung bereits bekannter heterologer GnTI-Gen-sequenzen mit hoher Wahrscheinlichkeit nicht erreicht werden kann.

Für Medizin und Forschung besteht daher nach wie vor ein Bedarf, rekombinante Glykoproteine mit minimalen und einheitlichen, also definierten Zuckerresten in geeigneten Organismen kostengünstig herstellen zu können.

### Wesen der Erfindung

Nachdem die Anmelderin erstmals pflanzliche GnTI-cDNA-Sequenzen isolieren und aufklären konnte, ist es nun u. a. möglich, beliebige Pflanzen mit gedrosselter oder fehlender GnTI-Aktivität zu gewinnen, insbesondere herzustellen, bzw. entsprechende Mutanten durch revers-genetische Ansätze nach Transposon-(Ref. 18) bzw. T-DNA-Insertion (Ref. 19) aufzuspüren, um in diesen dann Glykoproteine mit niedrigem Antigenpotential zu produzieren.

### 1) Enzyme

Die Erfindung umfaßt allgemein verschiedene N-Acetylglucosaminyltransferase I-Enzyme (EC 2.4.1.101) aus Pflanzen, z. B. aus Kartoffel (*Solanum tuberosum* L.), Tabak (*Nicotiana tabacum* L.) und Arabidopsis thaliana. Insbesondere betrifft die Erfindung die Enzyme, die die in Fig. 2 und 3B sowie in dem beigefügten Sequenzprotokoll angegebenen Aminosäuresequenzen aufweisen oder enthalten.

Von der Erfindung umfaßt sind ferner von den Aminosäuresequenzen der genannten Enzyme durch Aminosäuresubstitution, -deletion, -insertion, -modifikation oder durch C- und/oder N-terminale Verkürzung und/oder Verlängerung abgeleitete Enzyme, die, sofern sie enzymatische Aktivität zeigen, eine dem Ausgangsenzym vergleichbare Spezifität, also N-Acetylglucosaminyltransferase I-Aktivität, sowie ggf. vergleichbare Aktivität aufweisen.

Unter einer vergleichbaren Aktivität wird im vorliegenden Zusammenhang eine Aktivität verstanden, die bis zu 100% über oder unter der Aktivität des Ausgangsenzyms liegt. Von der Erfindung sind dementsprechend auch abgeleitete Enzyme oder Proteine mit sehr geringer oder vollständig fehlender enzymatischer Aktivität, nachweisbar mittels eines oder mehrerer der nachfolgend angegebenen Tests, umfaßt. Zum Nachweis der Enzymaktivität dient ein Standard-Test, der mit Mikrosomen-Fractionen entweder radioaktiv, z. B. mit UDP-[6-<sup>3</sup>H]GlcNAc als Substrat (Ref. 13), oder nicht-radioaktiv (HPLC-Methode; Ref. 20) durchgeführt wird. Pflanzliche GnTI-Aktivität kann subzellulär in Golgi-Fractionen nachgewiesen werden (Ref. 21). Enzymanreicherungen aus Pflanzen sind aufgrund der geringen Ausbeuten jedoch nahezu unmöglich.

Gegebenenfalls alternativ kann ein erfindungsgemäßes abgeleitetes Enzym als ein Enzym definiert werden, für das eine das Enzym kodierende DNA-Sequenz ermittelt oder abgeleitet werden kann, die mit der das Ausgangsenzym kodierenden DNA-Sequenz bzw. der dazu komplementären Sequenz unter stringenten Bedingungen, wie sie nachfolgend definiert werden, hybridisiert.

Ein dergestalt abgeleitetes Enzym stellt beispielsweise eine Isoform dar, die die Aminosäuren 74 bis 446 der in Fig. 2 und SEQ ID NO: 1 und 2 angegebenen Aminosäuresequenz umfaßt. Dieser Isoform fehlt u. a. der durch die Aminosäuren 10 bis 29 gebildete Membrananker, so daß diese Enzym-Isoform möglicherweise auf Grund dessen im Cytosol der Pflanze lokalisiert ist.

Als Beispiele für C- und/oder N-terminal verlängerte Proteine können Fusionsproteine genannt werden, die neben einer erfindungsgemäßen Aminosäuresequenz ein weiteres Protein umfassen, das beispielsweise eine andersartige enzymatische Aktivität aufweist oder auf andere Weise, z. B. aufgrund von Fluoreszenz oder Phosphoreszenz oder aufgrund einer Reaktivität mit spezifischen Antikörpern oder durch Binden an entsprechende Affinitätsmatrices, leicht nachweisbar ist.

Die Erfindung umfaßt gleichfalls Fragmente der genannten Enzyme, die ggf. keine enzymatische Aktivität mehr aufweisen. Diese Fragmente zeigen in der Regel jedoch antigene Wirkung in einem damit immunisierten Wirt und können folglich als Antigen zur Erzeugung monoklonaler oder polyklonaler Antikörper durch Immunisierung eines Wirtes mit diesen eingesetzt werden.

Die Erfindung betrifft darüberhinaus auch N-Acetylglucosaminyltransferase I-Enzyme insbesondere aus anderen Varietäten und Pflanzenarten, die zugänglich sind aufgrund der Hybridisierung ihrer Gene oder eines oder mehrerer Abschnitte ihrer Gene.

mit einer oder mehreren der nachfolgend erläuterten erfindungsgemäßen N-A-Sequenzen und/oder DNA-Fragmente und/oder  
 mit geeigneten erfindungsgemäßen Hybridisierungs sonden, die ausgehend von den im Sequenzprotokoll angegebenen Aminosäuresequenzen unter Berücksichtigung der Degeneration des genetischen Codes hergestellt werden können.

Von der Erfindung umfaßt sind ferner nach den vorstehend erläuterten Maßgaben von diesen N-Acetylglucosaminyltransferase I-Enzymen abgeleitete Enzyme oder Proteine, einschließlich Fusionsproteinen von diesen, sowie Fragmente aller dieser Enzyme oder Proteine.

## ii) Antikörper

Ein weiterer Aspekt der Erfindung betrifft die Verwendung der vorstehend genannten Aminosäuresequenzen bzw. von antigen wirksamen Fragmenten davon zur Erzeugung von monoklonalen oder polyklonalen Antikörpern oder -seren durch Immunisierung von Wirten mit diesen Aminosäuresequenzen bzw. Fragmenten sowie Antikörper bzw. -seren an sich, die die vorstehend erläuterten Enzyme und/oder Antigene spezifisch erkennen und binden. Die allgemeine Vorgehensweise und die entsprechenden Techniken zur Erzeugung polyklonaler und monoklonaler Antikörper sind dem Fachmann gleichfalls wohl bekannt.

Beispielhaft wurde rekombinantes GnIT-Protein aus *Solanum tuberosum* mit 10 N-terminalen Histidin-Resten ("His-tag") durch Verwendung eines Fragments der in **Fig. 2** und SEQ ID NO: 1 dargestellten GnIT-cDNA (Nukleotide 275 bis 1395) in *E. coli* überexprimiert und nach Affinitätsreinigung über eine Metall-Chelat-Matrix als Antigen zur Erzeugung von polyklonalen Antiseren in Kaninchen eingesetzt (vgl. Beispiele 5 und 6).

Eine Anwendungsmöglichkeit der erfindungsgemäßen Antikörper besteht für das "Screenen" von Pflanzen auf das Vorhandensein von N-Acetylglucosaminyltransferase I.

Eine Bindung des erfindungsgemäßen Antikörpers an pflanzliche(s) Proteine zeigt das Vorliegen eines mit diesem Antikörper nachweisbaren N-Acetylglucosaminyltransferase I-Enzyms an. In einem späteren Schritt kann dieser Antikörper darüberhinaus, in der Regel dann kovalent an einen Träger gebunden, gegebenenfalls auch zur Anreicherung oder Reinigung des Enzyms mittels Säulenchromatographie eingesetzt werden.

Ein negatives Bindungsergebnis mit dem erfindungsgemäßen Antikörper, d. h. fehlende Bindung an die pflanzlichen Proteine, läßt wiederum auf ein fehlendes (oder durch Mutation stark verändertes) N-Acetylglucosaminyltransferase I-Enzym und damit auf fehlende oder stark verminderte N-Acetylglucosaminyltransferase I-Aktivität in einer untersuchten Pflanze schließen.

Techniken zum Ausführen der vorstehend angesprochenen "Screening"-Tests oder zur Enzymanreicherung bzw. -reinigung unter Einsatz von Antikörpersäulen oder anderen Affinitätsmatrices (vgl. Beispiele 5 und 6) sind dem Fachmann wohl bekannt.

## iii) DNA-Sequenzen

Die Erfindung umfaßt ferner DNA-Sequenzen, die die erfindungsgemäßen Aminosäuresequenzen, einschließlich davon gemäß den vorstehenden Maßgaben abgeleitete Aminosäuresequenzen kodieren. Insbesondere betrifft die Erfindung das den in den **Fig. 2** und 3B und im Sequenzprotokoll aufgeführten Aminosäuresequenzen jeweils zugrundeliegende Gen, und ganz besonders die in **Fig. 2** und im Sequenzprotokoll aufgeführten cDNA-Sequenzen, sowie von diesen Genen und DNA-Sequenzen abgeleitete DNA-Sequenzen.

Unter abgeleiteten DNA-Sequenzen werden Sequenzen verstanden, die durch Substitution, Deletion und/oder Insertion von einzelnen oder mehreren und/oder kleineren Gruppen von Nukleotiden der vorstehend angegebenen Sequenzen und/oder durch Verkürzung oder Verlängerung am 5'- und/oder 3'-Ende erhalten werden. Die Modifizierungen innerhalb der DNA-Sequenz können zu abgeleiteten DNA-Sequenzen führen, die identische Aminosäuresequenzen verglichen mit der von der Ausgangs-DNA-Sequenz kodierten Aminosäuresequenz kodieren, oder aber auch zu solchen, bei denen einzelne oder einige wenige Aminosäuren gegenüber der Aminosäuresequenz, die die Ausgangs-DNA-Sequenz kodiert, verändert, d. h. substituiert, deletiert und/oder insertiert sind, oder auch solchen, die ggf. zusätzlich C- und/oder N-terminal verkürzt und/oder verlängert sind.

Die Erfindung erstreckt sich gleichfalls auf die zu den erfindungsgemäßen Genen und DNA-Sequenzen komplementären Sequenzen sowie auf deren RNA-Transkriptionsprodukte.

Von der Erfindung umfaßt sind insbesondere sämtliche, nach den vorstehend angegebenen Maßgaben abgeleiteten Sequenzen, die über ihre gesamte Länge hinweg oder lediglich mit einem oder mehreren Teilabschnitten unter stringenten Bedingungen mit den oben erläuterten Ausgangssequenzen oder den dazu komplementären Sequenzen oder Teilen davon hybridisieren, wie auch DNA-Sequenzen, die derartige Sequenzen umfassen.

Als Hybridisierung unter stringenten Bedingungen im Sinne der Erfindung wird eine Hybridisierung bei Vorgehensweise gemäß einem oder mehreren der nachstehend angegebenen Verfahren verstanden. Hybridisieren: Bis zu 20 Std. in PLG-Puffer nach Church und Gilbert (0,25 M  $\text{Na}_2\text{HPO}_4$ , 1 mM EDTA, 1% (w/v) BSA, 7% (w/v) SDS, pH 7,5 mit Phosphorsäure; Ref. 22) bei 42°C oder in Standard-Hybridisierungspuffern mit Formamid bei 42°C oder ohne Formamid bei 68°C (Ref. 23). Waschen: 3-mal 30 min bei 65°C in 3-fach SSC-Puffer (Ref. 23), 0,1% SDS.

Im Rahmen der vorliegenden Anmeldung ist der Begriff "Hybridisierung" stets als Hybridisierung unter stringenten Bedingungen, wie vorstehend angegeben, zu verstehen, auch wenn dies im Einzelfall nicht explizit angegeben ist.

Darüberhinaus erstreckt sich die Erfindung auch auf Fragmente der vorstehend erläuterten DNA-Sequenzen, einschließlich der nach den vorstehenden Maßgaben abgeleiteten DNA-Sequenzen, auf von derartigen Fragmenten durch Nukleinsäuresubstitution, -insertion und/oder -deletion abgeleitete Fragmente sowie auf die entsprechenden Fragmente mit dazu komplementären Sequenzen. Derartige Fragmente sind u. a. als Sequenzierungs- oder PCR-Primer, "Screening"

ning"-Sonden und/oder für Verwendungen, wie sie nachfolgend erläutert werden, geeignet. Für eine Verwendung als "Screening"- oder Hybridisierungssonde werden die erfindungsgemäßen DNA-Fragmente häufig radioaktiv markiert eingesetzt. Fragmente mit Sequenzen, die von den vorstehend definierten Ausgangssequenzen durch Substitution, Deletion und/oder Insertion von einem oder mehreren Nukleotiden abgeleitet sind, bzw. die dazu komplementären Sequenzen sind in dem Umfang von der Erfindung umfaßt, als sie unter den vorstehend angegebenen stringenten Bedingungen mit den Ausgangssequenzen bzw. den dazu komplementären Sequenzen hybridisieren.

Erfindungsgemäße DNA-Fragmente können beispielsweise auf der Grundlage der im Sequenzprotokoll und in Fig. 2 angegebenen DNA-Sequenzen ausgehend von pflanzlicher DNA mittels Restriktionsendonukleasen unter Nutzung geeigneter Restriktionsschnittstellen oder durch Einsatz von PCR mittels geeignet synthetisierter Primer erhalten oder alternativ auch chemisch synthetisiert werden. Derartige Techniken sind dem Fachmann wohlbekannt.

Die Erfindung betrifft darüberhinaus auch jegliche DNA-Sequenzen, die ein Gen darstellen oder Teil eines Gens sind, das das Enzym N-Acetylglucosaminyltransferase I kodiert, und die in ihrer Gesamtheit oder in einem Teilabschnitt

mit einer der erfindungsgemäßen DNA-Sequenzen und/oder  
mit einem oder mehreren der erfindungsgemäßen DNA-Fragmente und/oder  
mit einer DNA-Sequenz, die von den im Sequenzprotokoll angegebenen Aminosäuresequenzen unter Berücksichtigung der Degeneration des genetischen Codes abgeleitet worden ist, unter stringenten Bedingungen hybridisieren.

Als DNA-Fragmente werden hierzu Hybridisierungs- oder "Screening"-Sonden eingesetzt, die üblicherweise mindestens 15 Nukleotide, typischerweise zwischen 15 und 30 Nukleotide, gegebenenfalls aber auch wesentlich mehr, umfassen. Es können dafür beispielsweise die in Beispiel 1 eingesetzten Primer Verwendung finden. Alternativ können DNA-Sequenzen geeigneter Länge, abgeleitet von den im Sequenzprotokoll angegebenen DNA-Sequenzen, eingesetzt werden. Als dritte Möglichkeit können geeignete erfindungsgemäße Hybridisierungssonden ausgehend von den im Sequenzprotokoll angegebenen Aminosäuresequenzen unter Berücksichtigung der Degeneration des genetischen Codes entwickelt werden.

In diesem Sinne sind Gegenstand der Erfindung auch N-Acetylglucosaminyltransferase I kodierende Gene, die insbesondere aus anderen Varietäten oder Pflanzenarten aufgrund ihrer Hybridisierung mit den vorstehend angegebenen Hybridisierungssonden aufgefunden werden können, sowie davon gemäß den vorstehend erläuterten Maßgaben abgeleitete DNA-Sequenzen, DNA-Fragmente und -Konstrukte.

Die Isolierung des jeweiligen Gens und die Sequenzierung desselben nach der Auffindung mittels der erfindungsgemäßen Hybridisierungssonden liegen im Bereich der Fähigkeiten eines Fachmanns auf diesem Gebiet und sind exemplarisch in den Beispielen für N-Acetylglucosaminyltransferase I aus *Solanum tuberosum* und die entsprechenden Enzyme aus *Nicotiana tabacum* und *Arabidopsis thaliana* (Teilsequenz) erläutert.

Gegenstand der Erfindung sind schließlich auch "antisense"-Sequenzen bezüglich jeglichen vorstehend erläuterten DNA-Sequenzen.

#### iv) Konstrukte

Von der Erfindung werden auch Konstrukte umfaßt, die ggf. neben zusätzlichen 5'- und/oder 3'-Sequenzen, z. B. Linkern und/oder regulatorischen DNA-Sequenzen, oder andersartigen Modifikationen die erfindungsgemäßen DNA-Sequenzen, einschließlich der wie vorstehend ausgeführt abgeleiteten DNA-Sequenzen, umfassen.

Ein Beispiel hierfür sind Hybridisierungs- oder "Screening"-Sonden, die zusätzlich zu einer erfindungsgemäßen DNA-Sequenz noch ein in diesem Fall meist nichtradioaktives Nachweismittel für die Detektion von Hybridisierungsprodukten umfassen, z. B. fluoreszierende oder phosphoreszierende Moleküle, Biotin, Biotinderivate, Digoxigenin und Digoxigeninderivate. Es kommen in diesem Zusammenhang jedoch auch radioaktive oder nichtradioaktive Nachweismittel, die z. B. durch Endmarkierung an die erfindungsgemäße DNA-Sequenz angeheftet werden können, in Betracht.

Gegenstand der Erfindung sind auch "antisense"- und "sense"-Konstrukte bezüglich der erfindungsgemäßen DNA-Sequenzen und -Fragmente, nämlich bezüglich

der im Sequenzprotokoll angegebenen DNA-Sequenzen und der zugrundeliegenden Gene,  
der davon nach den vorstehenden Maßgaben abgeleiteten DNA-Sequenzen,  
eines oder mehrerer Abschnitte dieser DNA-Sequenzen,

DNA-Sequenzen insbesondere aus anderen Varietäten oder Pflanzenarten, die ein Gen darstellen oder Teil eines Gens sind, welches das Enzym N-Acetylglucosaminyltransferase I kodiert, und die  
mit einer der vorstehend genannten DNA-Sequenzen und/oder  
mit einem oder mehreren der vorstehend genannten DNA-Fragmente und/oder  
mit einer DNA-Sequenz, die von den im Sequenzprotokoll angegebenen Aminosäuresequenzen unter Berücksichtigung der Degeneration des genetischen Codes abgeleitet worden ist,  
unter stringenten Bedingungen hybridisieren.

Des weiteren erstreckt sich die Erfindung auch auf jegliche DNA-Übertragungssysteme, wie Vektoren, Plasmide, Viren- und Phagen Genome oder Cosmide, die die erfindungsgemäßen DNA-Sequenzen, z. B. das GntI-Gen, erfindungsgemäße cDNA und DNA-Abschnitte, wie sie im Sequenzprotokoll angegeben sind, Fragmente davon, insbesondere "antisense"- oder "sense"-Konstrukte und/oder davon nach den vorstehenden Maßgaben abgeleitete DNA-Sequenzen, enthalten.

Diverse Techniken zur Gewinnung oder Synthese erfindungsgemäßer DNA, DNA-Fragmente, Konstrukte und Übertragungssysteme, z. B. ausgehend von pflanzlicher DNA durch Restriktion mittels Restriktionsendonukleasen, PCR-

Amplifizierung unter Einsatz geeigneter Primer, ggf. gefolgt von Klonierung und optionaler chemischer oder enzymatischer Modifizierung, sind dem Fachmann wohlbekannt.

Eine Anwendungsmöglichkeit von erfindungsgemäßen DNA-Hybridisierungssonden liegt im Nachweis von N-Acetylglucosaminyltransferase I-Genen in anderen Pflanzen als jenen, aus denen die im Sequenzprotokoll angegebenen DNA-Sequenzen erhalten wurden, oder im Nachweis von möglichen (weiteren) Isoformen des N-Acetylglucosaminyltransferase I-Gens in den Ausgangspflanzen *Solanum tuberosum*, *Nicotiana tabacum* und *Arabidopsis thaliana*.

Kann für das Hybridisierungsexperiment auf eine genomische Bank oder cDNA-Bank einer Pflanze zurückgegriffen werden, liefert ein positives Hybridisierungsergebnis bei einem derartigen "Screening" oder Durchmustern der jeweiligen Bank den Hinweis auf einen Klon oder einige wenige Klone, die die gesuchte Sequenz, das N-Acetylglucosaminyltransferase I-Gen, vollständig oder teilweise in Verbindung mit nur einer begrenzten Menge weiterer DNA aus dem Genom der Zielpflanze enthalten, was die Klonierung und Sequenzierung des Zielgens entsprechend erleichtert. Alternativ kann ausgehend von pflanzlicher DNA und geeigneten Konstrukten, sogenannten PCR-Primern, auch eine PCR-Amplifikation des Gens oder von Teilen desselben vorgenommen werden, um Klonierung und Sequenzierung zu vereinfachen.

Ein Einsatz erfindungsgemäßer Sequenzierungsprimer, die ausgehend von geeigneten Abschnitten der erfindungsgemäßen Sequenzen synthetisiert werden, ermöglicht z. B. eine genomische Sequenzierung ausgehend von der vollständigen, durch Restriktionsendonukleasen geschnittenen genomischen DNA einer Zielpflanze mittels der Church-Gilbert-Technik wie auch z. B. eine Sequenzierung auf cDNA-Ebene nach RT-PCR-Amplifikation der Gesamt-RNA der Zielpflanze (vgl. Bsp. 1).

Eine alternative Anwendungsmöglichkeit von erfindungsgemäßen DNA-Hybridisierungssonden, die von den im Sequenzprotokoll angegebenen DNA-Sequenzen abgeleitet sind, besteht in der erfindungsgemäßen Verwendung derselben zum Nachweis von Pflanzen mit verminderter oder fehlender N-Acetylglucosaminyltransferase I-Aktivität. Das Hybridisierungsexperiment dient zur Detektion des Gens der N-Acetylglucosaminyltransferase I (*GntI*) und erlaubt z. B. aufgrund eines negativen Hybridisierungsergebnisses unter stringenten Bedingungen den Rückschluß auf ein Fehlen des *GntI*-Gens und damit fehlende N-Acetylglucosaminyltransferase I-Aktivität in einer untersuchten Pflanze.

Derartige Hybridisierungstechniken zum Nachweis von Proteinen oder Genen insbesondere in Pflanzenmaterial mittels DNA-Sonden sind dem Fachmann ebenfalls geläufig. Es wird in diesem Zusammenhang auf die vorstehenden Ausführungen zu möglichen Hybridisierungsbedingungen unter Punkt iii) verwiesen. Geeignete DNA-Hybridisierungssonden umfassen in der Regel mindestens 15 Nukleotide mit einer Sequenz, die z. B. von den in Fig. 2 und im Sequenzprotokoll angegebenen cDNA-Sequenzen oder den entsprechenden *GntI*-Genen abgeleitet ist.

#### v) Transformierte Mikroorganismen

Die Erfindung erstreckt sich des weiteren auf Mikroorganismen, wie z. B. Bakterien, Bakteriophagen, Viren, einzelne eukaryotische Organismen, wie Pilze, Hefen, Protozoen, Algen und humane, tierische und pflanzliche Zellen, die durch eine oder mehrere der erfindungsgemäßen DNA-Sequenzen oder eines oder mehrere der erfindungsgemäßen Konstrukte, wie vorstehend erläutert, transformiert wurden.

Verwendung finden erfindungsgemäße transformierte Mikroorganismen beispielsweise als Expressionssysteme für die transformierende Fremd-DNA zur Gewinnung der entsprechenden Expressionsprodukte. Typische Mikroorganismen für diese Zwecke sind Bakterien, wie beispielsweise *E. coli*. Des weiteren können erfindungsgemäße transformierte Mikroorganismen, insbesondere Agrobakterien, z. B. zur Transformation von Pflanzen unter Weitergabe der transformierenden Fremd-DNA eingesetzt werden.

Verfahren zur Transformation von Mikroorganismenzellen durch (Fremd-)DNA sind dem Fachmann wohlbekannt.

Hierfür werden z. B. als Expressionsvektoren bezeichnete Konstrukte eingesetzt, die die erfindungsgemäße DNA-Sequenz unter Kontrolle eines konstitutiven oder induzierbaren, ggf. zusätzlich gewebespezifischen Promotors enthalten, um eine Expression der eingeschleusten DNA in der Ziel- oder Wirtszelle zu ermöglichen.

Ein weiterer Aspekt der Erfindung ist dementsprechend ein Verfahren zur Gewinnung der erfindungsgemäßen Enzyme und Proteine unter Einsatz eines oder mehrerer der erfindungsgemäßen transformierten Mikroorganismen. Das Verfahren umfaßt, mindestens einen durch erfindungsgemäße DNA, insbesondere eine der im Sequenzprotokoll angegebenen cDNAs, unter Kontrolle eines aktiven Promotors transformierten Mikroorganismus, wie vorstehend definiert, zu züchten und das erfindungsgemäße Enzym aus den Mikroorganismen und gegebenenfalls auch dem Kulturmedium zu isolieren. Das Verfahren erstreckt sich selbstverständlich auch auf die Gewinnung von den erfindungsgemäßen Enzymen aus *Solanum tuberosum*, *Nicotiana tabacum* und *Arabidopsis thaliana* abgeleiteten Enzymen bzw. Proteinen, wie sie vorstehend unter i) definiert sind.

Verfahren zur Züchtung transformierter Mikroorganismen sind dem Fachmann wohlbekannt. Die Isolierung des exprimierten Enzyms kann beispielsweise gemäß dem in Beispiel 5 beschriebenen Verfahren mittels Metall-Chelat-Chromatographie oder alternativ durch Chromatographie an Säulen, die gegen das Enzym gerichtete Antikörper an das Packungsmaterial gebunden enthalten, erfolgen.

#### vi) Transgene Pflanzen

Die Erfindung umfaßt gleichfalls transgene Pflanzen, die mittels einer erfindungsgemäßen DNA-Sequenz bzw. eines entsprechenden Konstrukts transformiert worden sind. Es können so z. B. transgene Pflanzen erhalten werden, bei denen eine *GntI*-Defizienz, z. B. aufgrund eines fehlenden oder schadhafte *GntI*-Gens oder aufgrund von Defekten in den regulatorischen Bereichen dieses Gens, durch Komplementierung durch ein von den im Sequenzprotokoll angegebenen cDNA-Sequenzen abgeleitetes Konstrukt, dessen Expression unter Kontrolle eines aktiven konstitutiven oder induzierbaren, ggf. zusätzlich gewebespezifischen Promotors steht, beseitigt worden ist. Das aufgrund der in dem Konstrukt enthaltenen erfindungsgemäßen DNA exprimierte *GntI*-Enzym oder Protein mit *GntI*-Aktivität komplementiert in diesem Falle die in der Ausgangspflanze fehlende *GntI*-Aktivität.

Gleichfalls in Betracht gezogen sind transgene Pflanzen, in denen die in der Ausgangspflanze bereits vorhandene GntI-Aktivität durch zusätzliche Expression des durch ein erfindungsgemäßes Konstrukt eingeschleusten GntI-Transgens erhöht ist. Ein bei der Untersuchung des Enzyms N-Acetylglucosaminyltransferase I in Pflanzen bestehendes Hauptproblem war bislang die überaus geringe Expression des GntI-Gens in vivo, verbunden mit einer überaus geringen Enzymaktivität, die entsprechend schwer nachzuweisen war. Durch Coexpression einer erfindungsgemäßen DNA kann das Problem zu geringer GntI-Enzymaktivität bei Pflanzen behoben werden.

In diesem Falle kann es bevorzugt sein, für die Transformation von Pflanzen erfindungsgemäße DNA einzusetzen, die zusätzlich einen Sequenzabschnitt umfaßt, der nach Expression eine vereinfachte Detektierung und/oder Anreicherung bzw. Reinigung des Proteinproduktes mit GntI-Aktivität ermöglicht. Beispielsweise gelingt dies durch Einsatz einer speziellen DNA-Sequenz für die Expression eines rekombinanten GntI-Enzyms, die eine N- oder C-terminale Sequenzverlängerung trägt, die einen Affinitätsmarker kodiert. Ist zusätzlich ein Aminosäuresequenzabschnitt zwischen GntI-Enzym und Affinitätsmarker vorgesehen, der eine Erkennungsstelle für eine spezifische Protease darstellt, kann durch nachträglichen Einsatz dieser spezifischen Protease die N- oder C-terminale Sequenzverlängerung von dem GntI-Enzym abgespalten und das GntI-Enzym dadurch isoliert erhalten werden.

Ein Beispiel hierfür ist der Einsatz einer erfindungsgemäßen DNA-Sequenz, die das rekombinante GntI-Enzym mit einer C-terminalen Sequenzverlängerung, die den Affinitätsmarker AWRIPQFGG ("Strep-tag"; Ref. 39) kodiert, und einer dazwischenliegenden Protease-Erkennungsstelle, PEGR, kodiert. Die Expression der erfindungsgemäßen DNA liefert GntI-Enzyme mit der angegebenen C-terminalen Sequenzverlängerung, über welche die exprimierten Proteinmoleküle spezifisch an eine Streptavidin-derivatisierte Matrix binden und so isoliert werden können. Mittels der die Aminosäuresequenz PEGR spezifisch erkennenden Protease Faktor Xa kann der GntI-Anteil der Proteinmoleküle dann freigesetzt werden. Alternativ kann das vollständige Protein von der Streptavidin-derivatisierten Matrix mittels Biotin oder Biotinderivaten abgelöst werden.

Ein weiteres Beispiel stellen erfindungsgemäße DNA-Sequenzen dar, die ein Protein kodieren, das zusätzlich zu einem GntI-Enzym eine Mehrzahl, z. B. 10, N-terminal angefügte Histidin-Reste ("His-tag") umfaßt. Eine Isolierung bzw. Reinigung der exprimierten Proteine kann aufgrund der N-terminalen Histidin-Reste leicht durch Metall-Chelat-Affinitätschromatographie (z. B. Ni-Sephrose) erfolgen (vgl. auch Beispiel 5).

Die Erfindung umfaßt gleichfalls Teile derartiger transgener Pflanzen, entsprechend transformierte Pflanzenzellen, transgene Samen und transgenes Vermehrungsmaterial.

Ein weiterer zentraler Aspekt der Erfindung ist die Verwendung der vorstehend erläuterten Sequenzinformation zur Gewinnung von Pflanzen mit verminderter oder fehlender N-Acetylglucosaminyltransferase I-Aktivität.

Die Möglichkeiten zum Auffinden von Pflanzen mit verminderter oder fehlender N-Acetylglucosaminyltransferase I-Aktivität aufgrund eines Gendefektes oder fehlenden Gens durch Einsatz erfindungsgemäßer Antikörper oder erfindungsgemäßer "Screening"- oder Hybridisierungssonden wurden vorstehend bereits beschrieben.

Zwei weitere Möglichkeiten bestehen in der erfindungsgemäßen Verwendung von "antisense"- bzw. "sense"-DNA-Konstrukten, die von der DNA-Sequenz eines GntI-Gens einer Pflanze abgeleitet sind, zur Erzeugung transgener Pflanzen mit verminderter oder fehlender N-Acetylglucosaminyltransferase I-Aktivität mittels "homology-dependent gene silencing" (vgl. Ref. 16, 17). Die DNA-Sequenz, auf die zur Erzeugung der Konstrukte als Ausgangssequenz zurückgegriffen wird, kann dabei von der zu transformierenden Ausgangspflanze selbst oder aber auch von einer anderen Pflanzenvarietät oder -art stammen. Insbesondere finden "antisense"- oder "sense"-Konstrukte, wie sie vorstehend unter den Punkten iii) und iv) erläutert wurden, Verwendung. Die eingesetzten Konstrukte umfassen üblicherweise mindestens 50 bis 200 und mehr Basenpaare.

Insbesondere umfassen die hierfür eingesetzten Konstrukte mindestens 50 bis 200 und mehr Basenpaare mit einer Sequenz, die ausgehend

von den im Sequenzprotokoll angegebenen cDNA-Sequenzen und/oder den entsprechenden GntI-Genen und/oder

von den vorstehend erläuterten erfindungsgemäßen abgeleiteten DNA-Sequenzen und/oder DNA-Fragmenten und/oder

von DNA-Sequenzen insbesondere aus anderen Varietäten und Pflanzenarten, die N-Acetylglucosaminyltransferase I kodieren und aufgrund einer Hybridisierung mit Hybridisierungs- oder "Screening"-Sonden, wie sie vorstehend unter Punkt iii) und iv) definiert wurden, unter stringenten Bedingungen aufgefunden werden können,

abgeleitet wird.

Die Konstrukte enthalten in der Regel einen starken konstitutiven oder induzierbaren, ggf. zusätzlich gewebespezifischen Promotor, unter dessen Kontrolle die "antisense"- oder "sense"-DNA-Sequenzabschnitte stehen.

Bei der Erzeugung transgener Pflanzen durch Integration von "antisense"-Konstrukt(en) in das Pflanzengenom oder durch virale Infektion von Ausgangspflanzen oder -pflanzenzellen durch "antisense"-Konstrukt(e) enthaltendes Virus für eine extrachromosomale Propagation und Transkription des "antisense"-Konstrukts oder der "antisense"-Konstrukte in infiziertem Pflanzengewebe ist beabsichtigt, auf RNA-Ebene eine Hybridisierung von Transkripten des GntI-Gens mit Transkripten des "antisense"-DNA-Abschnitts zu erzielen, die die Translation der GntI-mRNA verhindert. Die Folge ist eine transgene Pflanze mit stark verringerten Gehalten an N-Acetylglucosaminyltransferase I und damit einer stark verringerten entsprechenden Enzymaktivität.

Für die erfindungsgemäße Transformation von Pflanzen mit "antisense"-Konstrukten können beispielsweise Konstrukte eingesetzt werden, die mit einer der kompletten, in Fig. 2 und im Sequenzprotokoll aufgeführten cDNAs oder entsprechenden, in der Regel mindestens 50 bis über 200 Basenpaare umfassenden Abschnitten derselben hybridisieren. Insbesondere bevorzugt ist darüberhinaus die Verwendung von Fragmenten, deren Transkripte zusätzlich zu einer Hybridisierung mit einem Teil des 5'-untranslatierten Bereiches der GntI-mRNA führen, an dem oder in dessen Nähe sich normalerweise die Anheftung der Ribosomen vollziehen würde. Beispiele für derartige Konstrukte sind in Fig. 4 gezeigt.

Angesichts des Vorkommens einer Isoform in *Solanum tuberosum* mit wahrscheinlich cytoplasmatischer Lokalisierung aufgrund des fehlenden Membranankers (As 10 bis 29) mit noch unbekannter Funktion kann es wünschenswert sein, lediglich das N-Acetylglucosaminyltransferase I-Enzym zu erfassen, das in den Golgi-Zisternen lokalisiert ist, d. h. nur jenes Enzym, das den Membrananker umfaßt. Ein Grund für diesen Wunsch kann das Bestreben oder im Einzelfalle auch die Notwendigkeit sein, den cytoplasmatischen Metabolismus der Pflanzenzelle, für den die cytoplasmatische N-Acetylglucosaminyltransferase I womöglich von Bedeutung ist, so wenig als möglich zu beeinträchtigen. Zu diesem Zweck können erfindungsgemäß "antisense"-Konstrukte eingesetzt werden, die bzw. deren Transkripte mit einem DNA- oder RNA-Abschnitt des GntI-Gens oder der GntI-mRNA hybridisieren, der einen Teil des 5'-untranslatierten Bereichs und den kodierenden Bereich einschließlich des Membranankers umfaßt. Eine Erstreckung des Hybridisierungsbereiches bis Position 266 der cDNA in **Fig. 2** und SEQ ID NO: 1 wird in der Regel für den genannten Zweck als unschädlich erachtet.

Bei der Erzeugung transgener Pflanzen durch Integration von "sense"-Konstrukten in das Pflanzengenom oder durch virale Infektion von Ausgangspflanzen oder -pflanzenzellen durch "sense"-Konstrukte(en) enthaltendes Virus für eine extrachromosomale Propagation und Expression des oder der Konstrukte in infiziertem Pflanzengewebe wird in Anlehnung an die Arbeiten von Fiske et al. (Ref. 17) in Tabak von Hybridisierungspänomenen dieser Konstrukte mit dem endogenen GntI-Gen auf posttranskriptionaler bzw. auf DNA-Ebene ausgegangen, die letztlich die Translation des GntI-Gens beeinträchtigen oder verhindern. Das Resultat sind auch in diesem Falle transgene Pflanzen mit verminderter oder sogar fehlender N-Acetylglucosaminyltransferase I-Aktivität.

Verfahren zur stabilen Integration derartiger "antisense"- und "sense"-Konstrukte in das Genom von Pflanzen bzw. zur viralen Infektion von Pflanzen bzw. Pflanzenzellen für eine extrachromosomale Propagation und Transkription/Expression derartiger Konstrukte in infiziertem Pflanzengewebe sind dem Fachmann bekannt. Dazu gehören sowohl der direkte DNA-Transfer (z. B. in Protoplasten mittels Elektroporation oder durch Zusatz eines hochmolekularen Osmotikums sowie biolistische Methoden, bei denen DNA-umhüllte Teilchen in Pflanzengewebe geschossen werden) wie die Verwendung natürlicher Wirt/Vektor-Systeme (z. B. Agrobakterien oder Pflanzenviren). Für eine virale Infektion von Ausgangspflanzen oder -pflanzenzellen durch entsprechende Konstrukte enthaltende Viren für eine extrachromosomale Propagation und Transkription/Expression der Konstrukte in infiziertem Pflanzengewebe stehen eine Reihe spezieller Viren, wie Tabakmosaikvirus (TMV) oder Kartoffelvirus X (potato virus X), zur Verfügung.

Beispielhafte Pflanzen, die für eine derartige Integration in Frage kommen, umfassen dikotyledone wie monokotyledone Kulturpflanzen, insbesondere Solanaceen wie Kartoffel, Tabak, Tomate und Paprika. Zusätzlich wären Banane, Luzerne, Raps, Rüben, Soja, Salat, Mais, Reis und Getreide geeignete Zielpflanzen für den Einsatz homologer "antisense"-Konstrukte. Beispielsweise erscheint die im Sequenzprotokoll angegebene Sequenz aus *Arabidopsis thaliana* insbesondere als Ausgangsequenz für die erfindungsgemäße Transformation von Brassicaceen, wie z. B. Rapspflanzen, mittels "sense"- oder "antisense"-Konstrukten geeignet. Weitere Pflanzen von Interesse sind jegliche Pflanzen, die für Medizin und Forschung interessante Glykoproteine exprimieren.

Allgemein soll festgehalten werden, daß die erfindungsgemäße Transformation von Pflanzen, die in dem entsprechenden Bereich des GntI-Gens eine Homologie von  $\geq 70\%$  auf Nukleotidebene zu den eingesetzten erfindungsgemäßen "antisense"- oder "sense"-Konstrukten aufweisen, in der Regel zu erfindungsgemäßen transgenen Pflanzen führt, die die gewünschte Verringerung der N-Acetylglucosaminyltransferase I-Aktivität aufweisen.

Eine weitere Möglichkeit wird ferner in einer zielgerichteten Zerstörung ("Knock out") des N-Acetylglucosaminyltransferase I-Gens durch "gene targeting" mittels homologer Rekombination (Ref. 24) in einer Zielpflanze durch Einsatz eines geeigneten, von der erfindungsgemäßen cDNA-Sequenz abgeleiteten DNA-Fragments gesehen, ähnlich der Vorgehensweise, wie sie beispielsweise für Hefe-Systeme und Säuger etabliert worden ist.

Die Erfindung umfaßt ferner transgene Pflanzen, die mit den vorstehend angesprochenen "antisense"- oder "sense"-Konstrukten bzw. mit diese enthaltenden Viren transformiert worden sind, sowie Teile derartiger transgener Pflanzen, entsprechend transformierte Pflanzenzellen, transgene Samen und transgenes Vermehrungsmaterial.

Verfahren zur Erzeugung transgener Pflanzen, z. B. durch Agrobakterien- oder Virus-vermittelten, wie auch direkten DNA-Transfer, sind dem Fachmann geläufig. Hinsichtlich beispielhafter Pflanzen für eine derartige Transformation gilt das vorstehend ausgeführte.

Die erfindungsgemäßen bzw. erfindungsgemäß gewonnenen Pflanzen mit verminderter oder fehlender N-Acetylglucosaminyltransferase I-Aktivität können erfindungsgemäß zur Herstellung von Glykoproteinen mit minimalen und einheitlichen, also definierten Zuckerresten eingesetzt werden. Wie bereits erläutert, sind derartige Glykoproteine für Medizin und Forschung von großer Bedeutung. Als preiswerte Rohstoff- und Nahrungsquelle sowie durch ihre problemlose Entsorgung über Kompostierung stellen Pflanzen per se ideale Bioreaktoren dar. Gemäß der vorstehend erläuterten Erfindung können jetzt biotechnologisch oder pharmazeutisch relevante Glykoproteine (z. B. Therapeutika mit niedrigem Antigenpotential für Säuger) in Kulturpflanzen exprimiert werden, in denen GntI-Aktivität stark gedrosselt ist oder vollständig fehlt.

Die Erfindung umfaßt dementsprechend auch ein Verfahren zur Erzeugung von Glykoproteinen mit minimalen, einheitlichen und definierten Zuckerresten, umfassend das Züchten einer erfindungsgemäßen transgenen Pflanze, von Teilen derartiger Pflanzen oder von erfindungsgemäßen transformierten Pflanzenzellen, die jeweils das gewünschte Glykoprotein exprimieren, und das Isolieren des gewünschten Glykoproteins aus dem gezüchteten Material.

Beispielhafte Kulturpflanzen sind in diesem Zusammenhang Solanaceen, insbesondere Kartoffel, Tabak, Tomate, und Paprika. Des weiteren kommen Banane, Luzerne, Raps, Rüben, Soja, Salat, Mais, Reis und Getreide in Frage.

Die Sequenz der enzymatisch gesteuerten, pflanzenspezifischen N-Glykan-Modifikationen, denen sekretorische Glykoproteine bei Passage durch den Golgi-Apparat höherer Pflanzen unterliegen, ist in **Fig. 1** schematisch dargestellt. Die Blockierung der Biosynthese durch fehlende oder unzureichende N-Acetylglucosaminyltransferase I-(GlcNAc-Transferase I)-Aktivität in einer Pflanze führt dazu, daß anstelle "komplexer" Glykane hauptsächlich Glykane des Man-GlcNAc-Typs, also Glykoproteine mit einheitlichen und wohldefinierten Zuckerresten gebildet werden, die für Medizin und Forschung von überaus hoher Bedeutung sind.



Hierzu können die Gene für die gewünschten Glykoproteine in ihren natürlichen Wirtspflanzen exprimiert werden, die erfindungsgemäß z. B. mittels "antisense"- oder "sense"-Konstrukten zu transgenen Pflanzen mit verminderter oder fehlender N-Acetylglucosaminyltransferase I-Aktivität transformiert wurden.

Es besteht jedoch auch die Möglichkeit, erfindungsgemäße transgene Pflanzen mit verminderter oder fehlender N-Acetylglucosaminyltransferase I-Aktivität einzusetzen, die zusätzlich mit dem Gen für das gesuchte Glykoprotein transformiert worden sind. Hierzu können Konstrukte eingesetzt werden, die das Gen für das gesuchte Glykoprotein unter der Kontrolle eines starken, konstitutiven oder induzierbaren, ggf. zusätzlich gewebespezifischen Promotors enthalten und die zu einer Integration des Gens in das Genom der Pflanze führen. Alternativ kann die Transformation auch durch virale Infektion durch ein das Gen für das gesuchte Glykoprotein enthaltendes Virus für eine extrachromosomale Propagation und Expression des Gens erfolgen. Das Glykoprotein kann dann in der jeweiligen Wirtspflanze exprimiert und daraus gewonnen werden.

Es kann dabei selbstverständlich alternativ auch so vorgegangen werden, daß zunächst eine Transformation mit einem Expressionskonstrukt oder Virus, das die kodierende DNA des Glykoproteins enthält, vorgenommen wird und erst danach eine weitere Transformation mit einem oder mehreren erfindungsgemäßen "antisense"- oder "sense"-Konstrukten oder einem oder mehreren Viren, die entsprechende DNA enthalten, erfolgt. Eine gleichzeitige Transformation mit beiden Konstrukten oder mit einem Virus, das sowohl das "antisense"- oder "sense"-Konstrukt als auch das das gesuchte Glykoprotein kodierende Gen enthält, ist ebenfalls möglich ("Tuckepack"-Version).

Im Rahmen der Erfindung wird auch eine virale Überinfektion erfindungsgemäßer transgener Pflanzen, bei denen bereits eine Integration des "antisense"/"sense"-Konstrukts und/oder des das gesuchte Glykoprotein kodierenden Gens im Genom vorliegt, durch Viren, die "antisense"/"sense"-Konstrukt und/oder das das gesuchte Glykoprotein kodierende Gen enthalten, für eine zusätzliche extrachromosomale Propagation und Transkription bzw. Expression dieser DNA in Betracht gezogen. Hierdurch können die Konzentrationen an "antisense"- bzw. "sense"-DNA oder exprimiertem Glykoprotein in den transgenen Pflanzenzellen erhöht werden.

Für die erfindungsgemäße Gewinnung definiert glykosylierter Glykoproteine kann sich eine Verwendung gewebespezifischer Promotoren beispielsweise in Fällen als sinnvoll erweisen, wenn beabsichtigt ist, die gewünschten Glykoproteine nur aus bestimmten Pflanzenteilen, wie der Knolle oder den Wurzeln, zu gewinnen. Für eine ganze Reihe von Pflanzengeweben stehen heute gewebespezifische Promotoren zur Verfügung, die eine Expression von Fremdgenen speziell nur in diesen Geweben bewirken. Beispielfähig können hier knollenspezifische Promotoren, wie Patatin Klasse I- (Ref. 26) und Proteinase Inhibitor II-Promotoren (Ref. 27) aufgeführt werden. Beide Promotoren zeigen unter bestimmten Bedingungen ebenfalls Expression in Blattgewebe, d. h. können durch hohe Metabolitgehalte (wie z. B. Saccharose) und im Fall des Proteinase Inhibitor II-Promotors auch durch mechanische Verwundung oder Besprühen mit Abscisin- bzw. Jasmonsäure induziert werden.

Eine Verwendung gewebespezifischer Promotoren kann auch dann angezeigt sein, wenn sich die für die Transformation eingesetzte erfindungsgemäße DNA-Sequenz bzw. deren Transkriptions- oder Translationsprodukte für bestimmte Pflanzenteile als abträglich erweisen, z. B. durch negative Einwirkung auf den Metabolismus der entsprechenden Pflanzenzellen.

Als beispielhaftes Ziel-Glykoprotein kommt humane Glucocerebrosidase zur Therapie der erblichen "Gaucher"-Krankheit (Ref. 25) in Frage. Zur Gewinnung von humaner Glucocerebrosidase (GC) mit einheitlichen und definierten Zuckerresten können beispielsweise erfindungsgemäße mittels "antisense"-DNA transformierte Pflanzen mit dem Gen für humane Glucocerebrosidase transformiert werden. Dafür wird die cDNA-Sequenz für humane Glucocerebrosidase (Ref. 38) mittels PCR unter Verwendung genspezifischer Primer am 3'-Ende so modifiziert, daß das rekombinante Enzym eine C-terminale Sequenzverlängerung trägt, die einen Affinitätsmarker (z. B. AWRIPQFGG, "Strep-tag"; Ref. 39) und gegebenenfalls auch eine Protease-Erkennungsstelle (z. B. PGIR) zwischen GntII-Enzymabschnitt und Affinitätsmarker kodiert. Die so veränderte GC-cDNA-Sequenz wird unter Verwendung eines starken und ggf. gewebespezifischen Promotors (z. B. für Kartoffel unter Kontrolle des knollenspezifischen B33-Patatin-Promotors) in erfindungsgemäßen GntI-antisense-Pflanzen exprimiert, so daß das in diesen Pflanzen synthetisierte Enzym ausschließlich wohldefinierte N-Glykane trägt. Der Affinitätsmarker soll die Anreicherung des rekombinanten Enzyms aus den transgenen Pflanzen erleichtern. In diesem Falle binden die exprimierten Proteinmoleküle ("GC-Strep"-Moleküle) über die Affinitätsmarkersequenz an eine Streptavidin-derivatisierte Matrix und können von dieser mittels Biotin oder Biotinderivaten abgelöst werden. Die Ablösung von der Streptavidin-derivatisierten Matrix kann auch mittels katalytischer Mengen einer Protease erfolgen, die eine Spezifität für die zwischen dem GntII-Enzymabschnitt und dem Affinitätsmarker befindliche Protease-Erkennungsstelle aufweist. In diesem Falle wird nur der GntII-Enzymabschnitt von der Matrix abgelöst. Dies kann insbesondere in dem Falle von Vorteil sein, wenn die Affinitätsmarkersequenz eine nachteilige Wirkung auf die GntII-Aktivität ausübt.

Die MannGlcNAc-Glykane der aus den erfindungsgemäßen Pflanzen gewonnenen Glucocerebrosidase werden aufgrund ihrer terminalen Mannose-Reste von Makrophagen als Aufnahmesignal erkannt und können daher direkt zur Therapie der erblichen "Gaucher-Krankheit" eingesetzt werden. Eine Therapie ist derzeit nur durch aufwendige Isolierung und Deglykosylierung nativer Glucocerebrosidase möglich (Ref. 25).

Die Herstellung rekombinanter Glykoproteine läßt sich dementsprechend durch den Einsatz pflanzlicher GntI-Sequenzen im Vergleich zu herkömmlichen Verfahren, z. B. der technisch aufwendigen chemischen Deglykosylierung gereinigter Glykoproteine (Ref. 25) oder einer schwierigen und teuren Produktion in GntII-defizienten tierischen Zelllinien (Ref. 7, 10), stark vereinfachen.

#### Erläuterung der Figuren

**Fig. 1** Sequenz der pflanzenspezifischen N-Glykan-Modifikationen, denen sekretorische Glykoproteine bei Passage durch den Golgi-Apparat höherer Pflanzen unterliegen (Ref. 28). Der Biosyntheseblock zu "komplex"-modifizierten Glykanen beruht auf einem Defizit an GntII-Aktivität (hervorgehoben entweder durch defektes oder fehlendes GntII-En-

zym oder durch effektive Diektierung der GntI-Genexpression) und ist durch ein Kreuz markiert. Bedeutung der Symbole: (F) Fucose-Reste, (X) Xylose-Reste, (●) GlcNac-Reste, (□) Mannose-Reste.

**Fig. 2** Vollständige cDNA-Sequenz einer pflanzlichen GntI aus Kartoffel (*Solanum tuberosum* L.) und davon abgeleitete Aminosäuresequenz. Exemplarisch ist die vollständige cDNA der GntI-Isoform mit Membrananker aus Kartoffelblattgewebe (A1) dargestellt. Die EcoRI/NotI-Linker an den 5'- und 3'-Enden der cDNA sind durch Fettdruck hervorgehoben, die Bindestellen der für die RT-PCR-Sonde verwendeten degenerierten Oligonukleotide sind unterstrichen. Im Gegensatz zu bereits publizierten tierischen GntI-Sequenzen enthält die abgeleitete Proteinsequenz der Kartoffel cDNA-Klone eine potentielle N-Glykosylierungsstelle: Asn-X(ohne Pro)-Ser/Thr, die mit einem Stern markiert ist. Die Region des Membranankers ist kursiv hervorgehoben (As 10 bis 29). Der Beginn der möglicherweise im Cytosol lokalisierten Isoform (A8) ist durch einen Pfeil gekennzeichnet.

**Fig. 3 A.** Identitäts- bzw. Ähnlichkeitsgrad der abgeleiteten Aminosäuresequenz einer vollständigen GntI-cDNA-Sequenz aus Kartoffel (A1) im Vergleich mit anderen, aus Datenbanken ausgewählten GntI-Sequenzen tierischer Organismen. Identische Aminosäurepositionen (in %) sind fett gedruckt, ähnliche Aminosäurepositionen stehen in Klammern darunter. Bedeutung der Kürzel: Hu, Mensch; Ra, Ratte; Mo, Maus; Ce, *Caenorhabditis elegans* (Spulwurm); St, *Solanum tuberosum* (Kartoffel).

**B.** Vergleich der abgeleiteten Aminosäuresequenzen verschiedener pflanzlicher GntI-cDNA-Klone. A Stb-A1, GntI aus Kartoffelblatt; B Ntb-A9, GntI aus Tabakblatt (A9); c Atb-Per, GntI-Teilsequenz aus *Arabidopsis thaliana* Blatt-mRNA (146 As). Identische As sind schwarz, ähnliche As hellgrau markiert.

**Fig. 4** Klonierungsschema der verwendeten GntI-"antisense"-Konstrukte. In die SalI-Schnittstelle der Polylinkerregion des pflanzlichen Expressionsvektors pA35 (Ref. 29) wurde nach Auffüllen der Enden ein NotI-Linker eingeführt (= pA35N) und die vollständige A1-GntI-cDNA über NotI in pA35N inseriert. Das entsprechende "antisense"-Konstrukt (= pA35N-A1as) wurde über EcoRI und HindIII in den binären Vektor pBin19 (Ref. 30) inseriert. Des weiteren wurde nach PCR-Amplifikation ein ca. 270 Bp umfassendes 5'-Fragment der A1-GntI-cDNA über XbaI- und NotI-Schnittstellen in pA35N in "antisense"-Orientierung kloniert (= pA35N-A1-kurz) und ebenfalls in pBin19 inseriert. Abkürzungen: Zahlen in Klammern, Positionsangaben der Restriktionsschnittstellen in der A1-GntI-cDNA (in Basenpaaren); pBSK, Klonierungsvektor (Stratagene); pGEM3Z, Klonierungsvektor (Promega); CaMV p35S, konstitutiver Blumenkohl-Mosaikvirus-35S-Promotor; OCSpA, Octopinsynthese-Polyadenylierungssignal; pNOS, Nopalinsynthese-Promotor; NI/O, Neomycinphosphotransferase (Selektionsmarker, vermittelt Kanamycinresistenz); NOSpA, Nopalinsynthese-Polyadenylierungssignal; LB/RB, "left/right border" der T-DNA des binären Vektors; Pfeil, Translationsstart (ATG); A8, Beginn der potentiell cytosolisch lokalisierten GntI-Isoform (7 As-Austausche im Vergleich zu A1).

#### Erläuterung der im Text verwendeten Abkürzungen

As, Aminosäure(n); Bp, Basenpaare; EMS, Ethylmethansulfonat (mutagene Chemikalie); F2, zweite Filialgeneration; Fuc, Fucose; Glc, Glucose; GlcNac, N-Acetylglucosamin; GntI, N-Acetylglucosaminyltransferase I (EC 2.4.1.101); GntI, Gen für GntI (Kern-codiert); Man, Mannose; PCR, Polymerasekettenreaktion; PAGE, Polyacrylamidgelelektrophorese; Ref., Referenz; RT-PCR, Reverse Transkription gekoppelt mit Polymerasekettenreaktion; SDS, Natriumdodecylsulfat; Var., Varietät; Xyl, Xylose.

Die Erfindung wird nun anhand von Beispielen eingehender erläutert. Die Beispiele werden lediglich zur Veranschaulichung der Erfindung aufgeführt und beschränken die Erfindung in keiner Weise.

Bsp. 1: Isolierung und Charakterisierung von pflanzlichen GntI-cDNA-Klonen

Aus Kartoffel- und Tabak-Blattgewebe wurde Gesamt-RNA isoliert und mittels RT-PCR in Kombination mit degenerierten Primern (Vorgehen analog zu Ref. 31), die von konservierten Aminosäurebereichen bekannter GntI-Sequenzen aus tierischen Organismen abgeleitet wurden ("sense"-Primer 1\*, 5'-TG(C/T) G(C/T)I (AT) (GC)T GCT TGG (AC)A(C/T) GA(C/T) AA(C/T)-3'; "antisense"-Primer 3\*, 5'-CCA ICC TT(AG) ICC (ACG)G(C/G) (AG)AA (AG)AA (AG)TC-3'; je 30 pMol Primer pro 50 µl PCR-Ansatz bei 55°C "annealing"-Temperatur und 45 Zyklen), cDNA-Fragmente von ca. 90 Bp amplifiziert. Nach Gelelektion wurden die Enden der PCR-Produkte repariert (d. h. durch DNA-Polymerase I glatte Enden erzeugt und mit T4-Polynukleotid-Kinase phosphoryliert) und in die EcoRV-Schnittstelle von pBSK (Stratagene) kloniert. Die Identität der abgeleiteten Aminosäuresequenzen der Kartoffel und Tabak RT-PCR-Produkte konnte durch Vergleich mit bekannten GntI-Sequenzen zwischen den Primern (Pfeile) als homolog bestätigt werden: => Q(R/M)QI VQDP(D/Y)AI YRS < (homologe As unterstrichen). Von je einem der Klone wurden mittels PCR radioaktiv markierte Sonden synthetisiert (Standard-PCR-Ansatz mit degenerierten Primern wie oben, Nukleotidmischung ohne dCTP, dafür mit 50 µCi α-<sup>32</sup>P-dCTP [ $>3000$  Ci/mMol] versetzt) und verschiedene cDNA-Banken mit den entsprechenden homologen Kartoffel- bzw. Tabak-Sonden auf GntI-haltige Klone durchgemustert (Vorgehen analog zu Ref. 31; "stringente" Hybridisierungsbedingungen wurden bereits oben im Text definiert). Die cDNA-Banken wurden mit mRNA aus jungen, noch wachsenden Pflanzenteilen ("sink"-Gewebe) hergestellt. Nach cDNA-Synthese und Ligieren von EcoRI/NotI-Adaptoren (cDNA-Synthese-Kit, Pharmacia) wurde mit EcoRI-kompatiblen Lambda-Armen ligiert, diese verpackt und damit E. coli XL1-Blue-Zellen transiziert (Lambda ZAPII Klonierungs- und Verpackungssystem, Stratagene). Nach Amplifikation der Banken wurde je ein vollständiger GntI-Klon aus einer Kartoffelblatt-"sink"-Bank (A1 gemäß **Fig. 2** und SEQ ID NO: 1) und einer Tabakblatt-"sink"-Bank (A9 gemäß SEQ ID NO: 3), sowie zwei weitere Klone aus einer Knollen"-"sink"-Bank isoliert (A6, A8). Die abgeleiteten GntI-Aminosäuresequenzen enthalten im Gegensatz zu denen von Tieren eine potentielle N-Glykosylierungsstelle, Asn-X(ohne Pro)-Ser/Thr. Eine der GntI-cDNA-Sequenzen aus Knollen trägt vor dem ersten Methionin Stop-Codons in allen drei Leserahmen (A8). Der codierende Bereich ist zu dem längeren Knollenklon (A6) stark homolog (nur 2 As-Austausche), trägt jedoch eine völlig andere 5'-untranslatierte Region. Des weiteren fehlt der für das Golgi-Enzym charakteristische Membrananker, so daß diese GntI-Isoform im Cytosol lokalisiert sein könnte. Sequenzvergleiche wurden mit Hilfe der "gap"- bzw. "pileup"- und "box"-Option des GCG-"software"-Pakets (J. Devereux, P. Haeblerli, O. Smithies (1984) Nucl Acids Res 12: 387-395) erstellt und zeigen, daß die abgeleiteten pflanzlichen GntI-Aminosäuresequenzen zu denen aus tierischen Organismen nur

30-40% Identität und 57-59% Ähnlichkeit aufweisen (**Fig. 3A**), untereinander aber hoch homolog sind ( $\geq 90\%$ , **Fig. 3B**).

Bei *Arabidopsis thaliana* wurde analog vorgegangen, wobei die erhaltene GntI-Teilsequenz durch RT-PCR mit GntI "sense"-Primer 4A (5'-ATCGGAAAGCTTCGGATCC CCA GTG GC(AG) GCT GTA GTT GTT ATG GCT TGC-3'; HindIII-Schnittstelle unterstrichen, BamHI fett gedruckt) und "antisense"-Primer 3\* wie vorstehend definiert, amplifiziert wurde. Die durch Sequenzierung ermittelte Nukleinsäuresequenz ist in SEQ ID NO: 5 aufgeführt.

Bsp. 2: Funktionelle Komplementierung eines GntI-Defekts mit GntI-cDNA bei transienter Expression in Protoplasten der *Arabidopsis thaliana* Cgl-Mutante.

Ca. 4 Wochen nach Aussaat wurden Protoplasten aus Blättern steril angezogener cgl-Mutanten ("Nichtfärber"-Pflanzen nach 5 Rückkreuzungen, Ref. 13) isoliert und mit Expressionskonstrukten der vollständigen GntI-cDNA-Sequenzen (NotI-cDNA-Fragmente, vgl. **Fig. 4**) in "sense"- (pA35N-A1s bzw. pA35N-A9s) oder "antisense"-Orientierung (pA35N-A1as bzw. pA35N-A9as) transformiert und für 96 Std. bei Raumtemperatur abgedunkelt kultiviert (je 50 µg Plasmid-DNA pro 1 Mio. Protoplasten, PEG-Methode nach Ref. 32). Anschließend SDS-PAGE der Protoplastenextrakte und "Western Blot"-Analyse (analog zu Ref. 13, 33) zeigte funktionelle Komplementierung des GntI-Defekts, d. h. "komplexe" Glykosylierung zahlreicher Proteinbanden bei transienter Expression der Kartoffel A1- und Tabak A9-"sense"-, nicht aber der entsprechenden "antisense"-Konstrukte in Protoplasten der *Arabidopsis* Cgl-Mutante (Daten nicht gezeigt).

Bsp. 3: Klonierung der binären Expressionskonstrukte pBin-35-A1as und pBin-35-A1-kurz (vgl. **Fig. 4**).

In die SalI-Schnittstelle der Polylinkerregion (entspricht puc18) des pflanzlichen Expressionsvektors pA35 (Ref. 29) wurde nach Auffüllen der Enden ein NotI-Linker eingeführt (pA35N), und die vollständige A1-GntI-cDNA (Nukleotide 9 bis 1657, gemäß der cDNA in **Fig. 2**) über NotI in pA35N inseriert ("sense"-Konstrukt pA35N-A1s bzw. "antisense"-Konstrukt pA35N-A1as). Die Expressionskassetten der "sense" bzw. "antisense"-Konstrukte wurden über die terminalen Schnittstellen (NotI-Schnittstelle aufgefüllt, mit HindIII partial nachverdaut) als ca. 2410 Bp-Fragment isoliert und in die EcoRI- (aufgefüllt) und HindIII-Schnittstellen des binären Vektors pBin19 (Ref. 30) inseriert (= pBin-35-A1s bzw. pBin-35-A1as). Durch Fusion mit der ebenfalls aufgefüllten NotI-Schnittstelle des Fragments wird die EcoRI-Schnittstelle des Vektors regeneriert. Zusätzlich wurde in einem Standard-PCR-Ansatz ("sense"-Primer: KS-Sequenzprimer (Stratagene), verlängert für PCR, 5'-GGC CCC CCC TCG AGG TCG ACG GTA TCG-3'; "antisense"-Primer: 5'-GGG C C T C T A G A C T C G A G A G C (C T) A C T A C T C T T C C T T G C T G C T G C T A A T C T T G-3', XbaI-Schnittstelle unterstrichen, XhoI-Schnittstelle kursiv) bei 50°C "annealing"-Temperatur ein 5'-Fragment der GntI-cDNA amplifiziert (Nukleotide 9 bis 261, gemäß der cDNA in **Fig. 2** und SEQ ID NO: 1). Das PCR-Produkt wurde mit XbaI (im "antisense"-Primer) und NotI (im 5'-Linker der cDNA) verdaut, als ca. 260 Bp-Fragment isoliert und in pA35N kloniert (= pA35N-A1-kurz). Die Expressionskassette des kurzen "antisense"-Konstrukts wurde als EcoRI/HindIII-Fragment (ca. 1020 Bp) ebenfalls in pBin19 inseriert (= pBin-35-A1-kurz).

Bsp. 4: Transformation von Agrobakterien durch die binären GntI-Konstrukte und Regeneration von transgenen Kartoffel- bzw. Tabakpflanzen aus infizierten Blattscheiben.

Die binären "antisense"-GntI-Konstrukte (pBin-35-A1as bzw. pBin-35-A1-kurz) wurden in den Agrobakterienstamm GV2260 transformiert (Ref. 34, 35). Mit den rekombinanten Agrobakterienlinien wurden exemplarisch sterile Blattscheiben von Kartoffelpflanzen der Var. Désirée bzw. Tabakpflanzen der Var. Wisconsin 38 infiziert (50 µl einer frischen Übernachtskultur in 10 ml flüssigem 2MS-Medium: 2% Saccharose in Murashige & Skoog Salz/Vitamin-Standard-Medium, pH 5,6; Blattstücke ohne Mittelrippen; Kokultivierung 2 Tage dunkel in Pflanzen-Klimakammer). Nach Waschen der infizierten Blattstücke in 2MS-Medium mit 250 µg/ml Clorfan wurden aus diesen in Gewebekultur unter Kanamycin-Selektion transgene Pflanzen regeneriert (Kartoffelprotokoll Ref. 26; Tabakprotokoll Ref. 36) und auf reduzierte GntI-Aktivität getestet (Daten nicht gezeigt).

Bsp. 5: Gewinnung von rekombinantem Kartoffel-GntI-Protein (zur Antikörperproduktion).

Mit Hilfe des pET-Systems (Novagen) wurde rekombinante GntI mit 10 zusätzlichen N-terminalen Histidin-Resten ("His-tag") in *E. coli* erzeugt und durch Metall-Chelat-Affinitätschromatographie gereinigt. Ein cDNA-Fragment, das die Nukleotide 275-1395 der Kartoffel GntI-cDNA umfaßt (entsp. As 75-446, **Fig. 2** bzw. SEQ ID NO: 1 und 2), wurde mittels Standard-PCR ("annealing"-Temperatur 50°C, 30 Zyklen, Ref. 31) amplifiziert ("sense"-Primer GntI-5'fus: 5'-CATGGATCC CTC GAG AAG CGT CAG GAC CAG GAG TGC CGG C-3'; "antisense"-Primer GntI-3'stop: 5'-ATCCCGGGATCCG CTA CGT ATC TTC AAC TCC AAG TTG-3'; XhoI- bzw. BamHI Schnittstellen unterstrichen, Stop-Codon kursiv), und über die Restriktionsschnittstellen der synthetischen Primer (5'-XhoI-GntI-BamHI-3') in den Vektor pET10b (Novagen) inseriert (= pET-His-A1). Nach Vermehrung und Analyse in *E. coli* XL1-Blue (Stratagene) wurde das Konstrukt als Glyzerinkultur archiviert. Zur Überexpression wurden kompetente *E. coli* BL21(DE3)pLysS-Zellen (Novagen) mit pET-His-A1 transformiert. Zugabe von IPTG (Isopropyl-1-thio-β-D-galaktopyranosid, ad 0,5-2 mM) zu einer logarithmisch wachsenden BL21-Kultur induziert zunächst die Expression von (bakterienchromosomaler) T7-RNA-Polymerase und damit ebenfalls die Expression des rekombinanten Fusionsproteins, das in pET-Vektoren (Novagen) unter T7-Promotor-Kontrolle steht. Aus induzierten BL21:pET-His-A1-Zellen wurde rekombinante Kartoffel-GntI mit "His-tag" unter denaturierenden Bedingungen (Hersteller-Protokoll, Novagen) mittels Metall-Chelat-Chromatographie an TALON-Matrix (Clontech) gereinigt und die Präparation mittels SDS-PAGE auf Einheitlichkeit überprüft.

Bsp. 6: Erzeugung von polyklonalen Antikörpern in Kaninchen.

Rekombinante Kartoffel-GntI (aus Bsp. 5) wurde als Antigen verwendet. Nach Entnahme von einigen Millilitern Prä-Immunserum wurde den Kaninchen in dreiwöchigen Abständen 300-500 µg affinitätsgereinigtes Protein zusammen mit 25 µg GMDP-Adjuvans (Gerbu) subcutan injiziert. Nach drei Basis-Injektionen wurden die Tiere 12 bis 14 Tage nach der jeweiligen Folge-Injektion ("Boost") aus der Ohrvene geblutet, das Serum gewonnen (Ref. 37) und auf Erkennung pflanzlicher GntI in "Western Blot"-Analysen (1:200 bis 1:2000-Verdünnung) getestet. Das Antiserum der "Boosts" mit geringstem Hintergrund-zu-Signal-Verhältnis wurde mit 0,04% Natriumazid versetzt, aliquotiert und bei +4°C bzw. längerfristig bei -20°C gelagert.

- 1) R Kornfeld, S Kornfeld (1985) Assembly of asparagine-linked oligosaccharides. *Annu Rev Biochem* 54: 631-664
- 2) GP Kaushal, T Szumilo, AD Elbein (1988). Structure and biosynthesis of plant N-linked glycans. In J Preiss (editor) *The Biochemistry of Plants*, Vol 14: Carbohydrates. Academic Press, San Diego, CA, pp 421-463
- 3) L Faye, MJ Chrispeels (1989) Apparent inhibition of  $\beta$ -fructosidase secretion by tunicamycin may be explained by breakdown of the unglycosylated protein during secretion. *Plant Physiol* 89: 845-851
- 4) TW Rademacher, RB Parekh, RA Dwek (1988) Glycobiology. *Annu Rev Biochem* 57: 785-838
- 5) A Sturm (1991) Heterogeneity of the complex N-linked oligosaccharides at specific glycosylation sites of 2 secreted carrot glycoproteins. *Eur J Biochem* 199: 169-179
- 6) K Olden, BA Bernard, MJ Humphries, T Yeo, SL White, SA Newton, HC Bower, JB Parent (1985) Function of glycoprotein glycans. *Trends Biochem Sci* 10: 78-82
- 7) P Stanley (1989) Chinese hamster ovary cell mutants with multiple glycosylation defects for production of glycoproteins with minimal carbohydrate heterogeneity. *Mol Cell Biol* 9: 377-383.
- 8) R Kumar, J Yang, RD Larsen, P Stanley (1990) Cloning and expression of N-acetylglucosaminyltransferase I, the medial Golgi transferase that initiates complex N-linked carbohydrate formation. *Proc Natl Acad Sci USA* 87: 9948-9952
- 9) JW Dennis, S Laferte, C Wagborne, ML Breitman, RS Kerbel (1987)  $\beta \rightarrow 6$  branching of Asn-linked oligosaccharides is directly associated with metastasis. *Science* 236: 582-585
- 10) MN Fukuda (1990) HbMPAS disease: genetic defect of glycosylation. *Glycobiology* 1: 9-15
- 11) MN Fukuda, KA Masri, A Dell, L Luzzatto, KW Moremen (1990) Incomplete synthesis of N-glycans in congenital dyserythropoietic anemia type II caused by a defect in the gene encoding  $\alpha$ -mannosidase II. *Proc Natl Acad Sci USA* 87: 7443-7447
- 12) M Lauriere, C Lauriere, MJ Chrispeels, KD Johnson, A Sturm (1989) Characterization of a xylose-specific antiserum that reacts with the complex asparagine-linked glycans of extracellular and vacuolar glycoproteins. *Plant Physiol* 90: 1182-1188
- 13) A von Schaewen, A Sturm, J O'Neill, MJ Chrispeels (1993) Isolation of a mutant Arabidopsis plant that lacks N-acetyl glucosaminyl transferase I and is unable to synthesize Golgi-modified complex N-linked glycans. *Plant Physiol* 102: 1109-1118
- 14) JK-C Ma, MB Hein (1995) Plant antibodies for immunotherapy. *Plant Physiol* 109: 341-346
- 15) AS Moffat (1995) Medical applications: Exploring transgenic plants as a new vaccine source. *Science* 268: 658-660 (Zusammenfassung von zwei Originalveröffentlichungen in derselben Ausgabe)
- 16) CB Taylor (1997) Comprehending cosuppression. *Plant Cell* 9: 1245-1249 (Zusammenfassung von mehreren Originalveröffentlichungen in derselben Ausgabe)
- 17) M Fiske, JE Backhausen, M Sendker, M Singer-Bayle, R Scheibe, A von Schaewen (1997) Transgenic tobacco plants expressing pea chloroplast NdhH cDNA in sense and antisense orientation: Effects on NADP-MDH level, stability of transformants, and plant growth. *Plant Physiol* 115: 705-715
- 18) R Koes, E Souer, A van Houwelingen, L Mur, C Spelt, F Quattrocchio, J Wing, B Oppedijk, S Ahmed, T Maes, T Gerats, P Hoogeveen, M Meesters, D Kloos, JNM Mol (1995) Targeted gene inactivation in petunia by PCR-based selection of transposon insertion mutants. *Proc Acad Sci USA* 92: 8149-8153
- 19) LC McKinney, N Ali, A Traut, KA Feldmann, DA Belostotsky, JM McDowell, RB Meagher (1995) Sequence-based identification of T-DNA insertion mutations in Arabidopsis: actin mutants act2-1 and act4-1. *Plant J* 8: 613-622
- 20) T Altmann, G Kornfeld, T Dalik, E Staudacher, J Glössl (1993) Processing of asparagine-linked oligosaccharides in insect cells. N-acetylglucosaminyl transferase I and II activities in cultured lepidopteran cells. *Glycobiology* 3: 619-625
- 21) A Sturm, KD Johnson, T Szumilo, AD Elbein, MJ Chrispeels (1987) Subcellular localization of glycosidases and glycosyltransferases involved in the processing of N-linked oligosaccharides. *Plant Physiol* 85: 741-745
- 22) GM Church, W Gilbert (1984) Genomic sequencing. *Proc Acad Sci USA* 81: 1991-1995
- 23) J Sambrook, EF Fritsch, T Maniatis (1989) Molecular cloning: a laboratory manual (2nd edn). Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- 24) H Puchta, B Hohn (1996) From centiMorgans to base pairs: homologous recombination in plants. *Plant Sci* 1: 340-348
- 25) NW Barton, FS Furbish, GJ Murray, M Gartfield, RO Brady (1990) Therapeutic response to intravenous infusions of glucocerebrosidase in a patient with Gaucher disease. *Proc Natl Acad Sci USA* 87: 1913-1916
- 26) M Rocha-Sosa, U Sonnewald, W-B Frommer, M Stratmann, J Schell, L Willmitzer (1989) Both developmental and metabolic signals activate the promoter of a class I patatin gene. *EMBO J* 8: 23-29
- 27) F Hildmann, M Ebner, H Pena-Cortes, JJ Sanchez-Serrano, L Willmitzer, S Prat (1992) General roles of abscisic and jasmonic acids in gene activation as a result of mechanical wounding. *Plant Cell* 4: 1157-1170
- 28) KD Johnson, MJ Chrispeels (1987) Substrate specificities of N-acetylglucosaminyl-, fucosyl-, and xylosyltransferases that modify glycoproteins in the Golgi apparatus of bean cotyledons. *Plant Physiol* 84: 1301-1308
- 29) H Hölte, L Faye, C Dickinson, EM Herman, MJ Chrispeels (1991) The protein-body proteins phytohemagglutinin and tonoplast intrinsic protein are targeted to vacuoles in leaves of transgenic tobacco. *Planta* 184: 431-437
- 30) M Bevan (1984) Binary Agrobacterium vectors for plant transformation. *Nucl Acids Res* 12: 8711-8721
- 31) K Graeve, A von Schaewen, R Scheibe (1994) Purification, characterization and cDNA sequence of glucose-6-phosphate dehydrogenase from potato (*Solanum tuberosum* L.). *Plant J* 5: 353-361
- 32) B Danini, R Schmidt, L Willmitzer (1989) Efficient transformation of Arabidopsis thaliana using direct gene

transfer to protoplasts. *Mol Cell Biochem* 213: 15-20

33) A von Schaewen, M Stitt, R Schmidt, L Willmitzer (1990) Expression of a yeast-derived invertase in the cell wall of tobacco and Arabidopsis plants leads to accumulation of carbohydrate, inhibition of photosynthesis and strongly influences growth and phenotype of transgenic tobacco plants. *EMBO J* 9: 3033-3044

34) R Deblaere, B Bytebier, H De Greve, F Debrouck, J Schell, M van Montagu, J Leemans (1985) Efficient octopine Ti plasmid-derived vectors for Agrobacterium mediated gene transfer to plants. *Nucl Acids Res* 13: 4777-4788

35) R Höfgen, L Willmitzer (1988) Storage of competent cells for Agrobacterium transformation. *Nucl Acids Res* 16: 9877

36) T Voelker, A Sturm, MJ Chrispeels (1987) Differences in expression between two seed lectin alleles obtained from normal and lectin-deficient beans are maintained in transgenic tobacco. *EMBO J* 6: 3571-3577

37) E Harlow, D Lane (1988) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

38) J Sorge, C West, B Westwood, E Beutler (1985) Molecular cloning and nucleotide sequence of human cerebrosidase cDNA. *Proc Natl Acad Sci USA* 82: 7289-7293

39) TGM Schmidt, A Skerra (1993) The random peptide library-assisted engineering of a C-terminal affinity peptide, useful for the detection and purification of a functional Ig Fv fragment. *Prot Engineering* 6: 109-122.

## (1) ALLGEMEINE ANGABEN:

## (i) ANMELDER:

- (A) NAME: von Schaewen, Antje, Dr. rer. nat.
- (B) STRASSE: Natruperstrasse 169a
- (C) ORT: Osnabrück
- (E) LAND: Deutschland
- (F) POSTLEITZAHL: 49076
- (G) TELEFON: 0541-684029

(ii) BEZEICHNUNG DER ERFINDUNG: Herstellung von Glykoproteinen in Pflanzen mit verminderter oder fehlender N-Acetylglucosaminyltransferase I (GnTI)-Aktivitaet unter Verwendung pflanzlicher gntI-Sequenzen

(iii) ANZAHL DER SEQUENZEN: 6

## (iv) COMPUTER-LESBARE FASSUNG:

- (A) DATENTRÄGER: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) BETRIEBSSYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPA)

## (2) ANGABEN ZU SEQ ID NO: 1:

## (i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1669 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Doppelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNA zu mRNA

(iii) HYPOTHETISCH: NEIN

(iv) ANTISENSE: NEIN

## (vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Solanum tuberosum
- (B) STAMM: Desiree
- (D) ENTWICKLUNGSSTADIUM: "sink" Organ
- (F) GEWEBETYP: Mesophyll
- (G) ZELLTYP: Blattzellen

## (vii) UNMITTELBARE HERKUNFT:

- (A) BIBLIOTHEK: Lambda ZAP II (EcoRI)
- (B) CLON(E): gntI-A1(K)

5

## (ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: misc\_feature
- (B) LAGE:659..667
- (D) SONSTIGE ANGABEN:/function= "Asn-Codon ist eine  
potentielle Glykosylierungsstelle"  
/product= "Konsensus Sequenz fuer  
N-Glykosylierung"  
/phenotype= "N-Glykane modulieren  
Proteineigenschaften"  
/standard\_name= "N-Glykosylierungsstelle"  
/label= pot-CHO  
/note= "fehlt in tierischen GntI-Sequenzen"

10

15

20

## (ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: CDS
- (B) LAGE:53..1393
- (C) ART DER ERMITTLUNG: experimentell
- (D) SONSTIGE ANGABEN:/codon\_start= 53  
/function= "initiiert komplexe N-Glykane auf  
sekret. Glykoproteinen"  
/EC\_number= 2.4.1.101  
/product=  
"beta-1,2-N-Acetylglucosaminyltransferase I"  
/evidence= EXPERIMENTAL  
/gene= "cgl"  
/standard\_name= "gntI"  
/label= ORF  
/note= "erste gntI-Sequenz aus Kartoffel  
(unpubliziert)"

25

30

35

40

45

## (ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: 5'UTR
- (B) LAGE:15..52

50

## (ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: 3'UTR
- (B) LAGE:1394..1655

55

## (ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: CDS
- (B) LAGE:80..139
- (D) SONSTIGE ANGABEN:/function= "Membrananker  
(Aminosäuren 10-29)"  
/product= "hydrophober Aminosäurebereich in  
GntI"

60

65

/standard\_name= "Membrananker eines Typ II  
Golgi-Proteins"  
/note= "durch Vergleich mit tierischen GnTI-  
Sequenzen identifiziert"

## (ix) MERKMAL:

(A) NAME/SCHLÜSSEL: misc\_feature  
(B) LAGE:1..14  
(D) SONSTIGE ANGABEN:/function= "zur Herstellung der  
cDNA-Bank in Lambda ZAP II"  
/product= "EcoRI/NotI-cDNA-Adaptor"  
/number= 1

## (ix) MERKMAL:

(A) NAME/SCHLÜSSEL: misc\_feature  
(B) LAGE:1656..1669  
(D) SONSTIGE ANGABEN:/product= "EcoRI/NotI-Adaptor"  
/number= 2

## (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 1:

GAATTCGCGG CCGCCTGAGA AACCTCGAA TTCAATTTCTG CATTGGCAG AG ATG	55
Met	
1	
AGA GGG AAC AAG TTT TGC TTT GAT TTA CGG TAC CTT CTC GTC GTG GCT	103
Arg Gly Asn Lys Phe Cys Phe Asp Leu Arg Tyr Leu Leu Val Val Ala	
5 10 15	
GCT CTC GCC TTC ATC TAC ATA CAG ATG CGG CTT TTC GCG ACA CAG TCA	151
Ala Leu Ala Phe Ile Tyr Ile Gln Met Arg Leu Phe Ala Thr Gln Ser	
20 25 30	
GAA TAT GTA GAC CGC CTT GCT GCT GCA ATT GAA GCA GAA AAT CAT TGT	199
Glu Tyr Val Asp Arg Leu Ala Ala Ala Ile Glu Ala Glu Asn His Cys	
35 40 45	
ACA AGT CAG ACC AGA TTG CTT ATT GAC AAG ATT AGC CAG CAG CAA GGA	247
Thr Ser Gln Thr Arg Leu Leu Ile Asp Lys Ile Ser Gln Gln Gln Gly	
50 55 60 65	
AGA GTA GTA GCT CTT GAA GAA CAA ATG AAG CAT CAG GAC CAG GAG TGC	295
Arg Val Val Ala Leu Glu Glu Gln Met Lys His Gln Asp Gln Glu Cys	
70 75 80	
CGG CAA TTA AGG GCT CTT GTT CAG GAT CTT GAA AGT AAG GGC ATA AAA	343
Arg Gln Leu Arg Ala Leu Val Gln Asp Leu Glu Ser Lys Gly Ile Lys	
85 90 95	
AAG TTA ATC GGA GAT GTG CAG ATG CCA GTG GCA GCT GTA GTT GTT ATG	391
Lys Leu Ile Gly Asp Val Gln Met Pro Val Ala Ala Val Val Val Met	
100 105 110	



# DE 197 54 622 A 1

GCT TGC AGT CGT ACT GAC TAC CTG GAG AGG ACT ATT AAA TCC ATC TTA	439	
Ala Cys Ser Arg Thr Asp Tyr Leu Glu Arg Thr Ile Lys Ser Ile Leu		
115 120 125		
AAA TAC CAA ACA TCT GTT GCA TCA AAA TAT CCT CTT TTC ATA TCC CAG	487	
Lys Tyr Gln Thr Ser Val Ala Ser Lys Tyr Pro Leu Phe Ile Ser Gln		
130 135 140 145		
GAT GGA TCA AAT CCT GAT GTA AGA AAG CTT GCT TTG AGC TAT GGT CAG	535	
Asp Gly Ser Asn Pro Asp Val Arg Lys Leu Ala Leu Ser Tyr Gly Gln		
150 155 160		
CTG ACG TAT ATG CAG CAC TTG GAT TAT GAA CCT GTG CAT ACT GAA AGA	583	
Leu Thr Tyr Met Gln His Leu Asp Tyr Glu Pro Val His Thr Glu Arg		
165 170 175		
CCA GGG GAA CTG GTT GCA TAC TAC AAG ATT GCA CGT CAT TAC AAG TGG	631	
Pro Gly Glu Leu Val Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys Trp		
180 185 190		
GCA TTG GAT CAG CTG TTT CAC AAG CAT AAT TTT AGC CGT GTT ATC ATA	679	
Ala Leu Asp Gln Leu Phe His Lys His Asn Phe Ser Arg Val Ile Ile		
195 200 205		
CTA GAA GAT GAT ATG GAA ATT GCT GCT GAT TTT TTT GAC TAT TTT GAG	727	
Leu Glu Asp Asp Met Glu Ile Ala Ala Asp Phe Phe Asp Tyr Phe Glu		
210 215 220 225		
GCT GGA GCT ACT CTT CTT GAC AGA GAC AAG TCG ATT ATG GCT ATT TCT	775	
Ala Gly Ala Thr Leu Leu Asp Arg Asp Lys Ser Ile Met Ala Ile Ser		
230 235 240		
TCT TGG AAT GAC AAT GGA CAA AGG CAG TTC GTC CAA GAT CCT GAT GCT	823	
Ser Trp Asn Asp Asn Gly Gln Arg Gln Phe Val Gln Asp Pro Asp Ala		
245 250 255		
CTT TAC CGC TCA GAC TTT TTT CCT GGT CTT GGA TGG ATG CTT TCA AAA	871	
Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Ser Lys		
260 265 270		
TCA ACT TGG TCC GAA CTA TCT CCA AAG TGG CCA AAG GCT TAC TGG GAT	919	
Ser Thr Trp Ser Glu Leu Ser Pro Lys Trp Pro Lys Ala Tyr Trp Asp		
275 280 285		
GAC TGG CTA AGG CTG AAA GAA AAT CAC AGA GGT CGA CAA TTT ATT CGC	967	
Asp Trp Leu Arg Leu Lys Glu Asn His Arg Gly Arg Gln Phe Ile Arg		
290 295 300 305		
CCA GAA GTT TGC AGA ACG TAC AAT TTT GGT GAG CAT GGT TCT AGT TTG	1015	
Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly Ser Ser Leu		
310 315 320		
GGG CAG TTT TTT AAG CAG TAT CTT GAG CCA ATT AAG CTA AAT GAT GTC	1063	
Gly Gln Phe Phe Lys Gln Tyr Leu Glu Pro Ile Lys Leu Asn Asp Val		
325 330 335		
CAG GTT GAT TGG AAG TCA ATG GAC CTA AGT TAC CTT TTG GAG GAC AAC	1111	
Gln Val Asp Trp Lys Ser Met Asp Leu Ser Tyr Leu Leu Glu Asp Asn		
340 345 350		

# DE 197 54 622 A 1

TAT GTG AAA CAC TTT GGC GAC TTG GTT AAA AAG GCT AAG CCC ATC CAC 1159  
Tyr Val Lys His Phe Gly Asp Leu Val Lys Lys Ala Lys Pro Ile His  
355 360 365

GGA GCT GAT GCT GTT TTG AAA GCA TTT AAC ATA GAT GGT GAT GTG CGT 1207  
Gly Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly Asp Val Arg  
370 375 380 385

ATT CAG TAC AGA GAC CAA CTA GAC TTT GAA GAT ATC GCT CGA CAG TTT 1255  
Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asp Ile Ala Arg Gln Phe  
390 395 400

GGC ATT TTT GAA GAA TGG AAG GAT GGT GTA CCA CGG GCA GCA TAT AAA 1303  
Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala Ala Tyr Lys  
405 410 415

GGG ATA GTA GTT TTC CGG TTT CAA ACA TCT AGA CGT GTG TTC CTT GTT 1351  
Gly Ile Val Val Phe Arg Phe Gln Thr Ser Arg Arg Val Phe Leu Val  
420 425 430

TCC CCT GAT TCT CTT CGA CAA CTT GGA GTT GAA GAT ACT TAG 1393  
Ser Pro Asp Ser Leu Arg Gln Leu Gly Val Glu Asp Thr \*  
435 440 445

CGAAGATATG ATTGGAGCCT GAGCAACAAT TTAGACTTAT TTGGTAGGAT ACATTTGAAA 1453

GAGCTGACAC GAAAAGTATG ACTACCAGTA GCTACATGCA ACATTTTAAT GTTAATGGAA 1513

GGAACCCACT GCTTATTGTT GGAATGGATG AATCATCACC ACATCCTATT ATTCAAGTTT 1573

ACAAACATAA AGAGGAAATG TTGCCCTATA AAAACAAATT TTTTGTTTCT AAGAAGGAAC 1633

GTTACGATTA TGAGCAACTT TGGCGGCCGC GAATTC 1669

## (2) ANGABEN ZU SEQ ID NO: 2:

### (i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 447 Aminosäuren
- (B) ART: Aminosäure
- (D) TOPOLOGIE: linear

### (ii) ART DES MOLEKÜLS: Protein

### (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 2:

Met Arg Gly Asn Lys Phe Cys Phe Asp Leu Arg Tyr Leu Leu Val Val  
1 5 10 15

Ala Ala Leu Ala Phe Ile Tyr Ile Gln Met Arg Leu Phe Ala Thr Gln  
20 25 30

Ser Glu Tyr Val Asp Arg Leu Ala Ala Ile Glu Ala Glu Asn His  
35 40 45

Cys Thr Ser Gln Thr Arg Leu Leu Ile Asp Lys Ile Ser Gln Gln Gln  
50 55 60

## DE 197 54 622 A 1

Gly	Arg	Val	Val	Ala	Leu	Glu	Glu	Gln	Met	Lys	His	Gln	Asp	Gln	Glu			
65					70					75					80			
Cys	Arg	Gln	Leu	Arg	Ala	Leu	Val	Gln	Asp	Leu	Glu	Ser	Lys	Gly	Ile			5
				85					90					95				
Lys	Lys	Leu	Ile	Gly	Asp	Val	Gln	Met	Pro	Val	Ala	Ala	Val	Val	Val			10
			100					105					110					
Met	Ala	Cys	Ser	Arg	Thr	Asp	Tyr	Leu	Glu	Arg	Thr	Ile	Lys	Ser	Ile			
		115					120					125						
Leu	Lys	Tyr	Gln	Thr	Ser	Val	Ala	Ser	Lys	Tyr	Pro	Leu	Phe	Ile	Ser			15
		130				135					140							
Gln	Asp	Gly	Ser	Asn	Pro	Asp	Val	Arg	Lys	Leu	Ala	Leu	Ser	Tyr	Gly			20
145					150					155					160			
Gln	Leu	Thr	Tyr	Met	Gln	His	Leu	Asp	Tyr	Glu	Pro	Val	His	Thr	Glu			
				165					170					175				
Arg	Pro	Gly	Glu	Leu	Val	Ala	Tyr	Tyr	Lys	Ile	Ala	Arg	His	Tyr	Lys			25
			180					185					190					
Trp	Ala	Leu	Asp	Gln	Leu	Phe	His	Lys	His	Asn	Phe	Ser	Arg	Val	Ile			30
		195					200					205						
Ile	Leu	Glu	Asp	Asp	Met	Glu	Ile	Ala	Ala	Asp	Phe	Phe	Asp	Tyr	Phe			
		210				215					220							
Glu	Ala	Gly	Ala	Thr	Leu	Leu	Asp	Arg	Asp	Lys	Ser	Ile	Met	Ala	Ile			35
225					230					235				240				
Ser	Ser	Trp	Asn	Asp	Asn	Gly	Gln	Arg	Gln	Phe	Val	Gln	Asp	Pro	Asp			40
			245					250					255					
Ala	Leu	Tyr	Arg	Ser	Asp	Phe	Phe	Pro	Gly	Leu	Gly	Trp	Met	Leu	Ser			
			260					265					270					
Lys	Ser	Thr	Trp	Ser	Glu	Leu	Ser	Pro	Lys	Trp	Pro	Lys	Ala	Tyr	Trp			45
			275				280					285						
Asp	Asp	Trp	Leu	Arg	Leu	Lys	Glu	Asn	His	Arg	Gly	Arg	Gln	Phe	Ile			50
		290				295					300							
Arg	Pro	Glu	Val	Cys	Arg	Thr	Tyr	Asn	Phe	Gly	Glu	His	Gly	Ser	Ser			
305					310					315				320				
Leu	Gly	Gln	Phe	Phe	Lys	Gln	Tyr	Leu	Glu	Pro	Ile	Lys	Leu	Asn	Asp			55
				325				330					335					
Val	Gln	Val	Asp	Trp	Lys	Ser	Met	Asp	Leu	Ser	Tyr	Leu	Leu	Glu	Asp			60
			340					345					350					
Asn	Tyr	Val	Lys	His	Phe	Gly	Asp	Leu	Val	Lys	Lys	Ala	Lys	Pro	Ile			
		355					360					365						65

His Gly Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly Asp Val  
 370 375 380

5 Arg Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asp Ile Ala Arg Gln  
 385 390 395 400

10 Phe Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala Ala Tyr  
 405 410 415

Lys Gly Ile Val Val Phe Arg Phe Gln Thr Ser Arg Arg Val Phe Leu  
 420 425 430

15 Val Ser Pro Asp Ser Leu Arg Gln Leu Gly Val Glu Asp Thr \*  
 435 440 445

(2) ANGABEN ZU SEQ ID NO: 3:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 1737 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Doppelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNA zu mRNA

(iii) HYPOTHETISCH: NEIN

(iv) ANTISENSE: NEIN

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Nicotiana tabacum
- (B) STAMM: Samsun NN
- (D) ENTWICKLUNGSSTADIUM: "sink" Organ
- (F) GEWEBETYP: Mesophyll
- (G) ZELLTYP: Blattzellen

(vii) UNMITTELBARE HERKUNFT:

- (A) BIBLIOTHEK: Lambda ZAP II (EcoRI)
- (B) CLON(E): gntI-A9(T)

(ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: misc\_feature
- (B) LAGE: 733..741
- (D) SONSTIGE ANGABEN: /function= "Asn-Codon ist eine  
 potentielle Glykosylierungsstelle"  
 /product= "Konsensus Sequenz für  
 N-Glykosylierung"  
 /phenotype= "N-Glykane modulieren  
 Proteineigenschaften"  
 /standard\_name= "N-Glykosylierungsstelle"  
 /label= pot-CHO  
 /note= "fehlt in tierischen GntI-Sequenzen"

## (ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: CDS
- (B) LAGE:127..1467
- (C) ART DER ERMITTLUNG: experimentell
- (D) SONSTIGE ANGABEN:/codon\_start= 127  
/function= "initiiert komplexe N-Glykane auf  
sekret. Glykoproteinen"  
/EC\_number= 2.4.1.101  
/product=  
"beta-1,2-N-Acetylglucosaminyltransferase I"  
/evidence= EXPERIMENTAL  
/gene= "cgl"  
/standard\_name= "gntI"  
/label= ORF  
/note= "erste gntI-Sequenz aus Tabak  
(unpubliziert)"

## (ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: 5'UTR
- (B) LAGE:15..126

## (ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: 3'UTR
- (B) LAGE:1468..1723

## (ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: CDS
- (B) LAGE:154..213
- (D) SONSTIGE ANGABEN:/function= "Membrananker  
(Aminosäuren 10-29)"  
/product= "hydrophober Aminosäurebereich in  
GnTI"  
/standard\_name= "Membrananker eines Typ II  
Golgi-Proteins"  
/note= "durch Vergleich mit tierischen GnTI-  
Sequenzen identifiziert"

## (ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: misc\_feature
- (B) LAGE:1..14
- (D) SONSTIGE ANGABEN:/function= "zur Herstellung der  
cDNA-Bank in Lambda ZAP II"  
/product= "EcoRI/NotI-cDNA-Adaptor"  
/number= 1

## (ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: misc\_feature
- (B) LAGE:1724..1737

(D) SONSTIGE ANGABEN:/product= "EcoRI/NotI-Adaptor"  
/number= 2

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 3:

10	GAATTCGCGG CCGCCATTGA CTTGATCCTA ACTGAACAGG CAAAGTAAAT CCAGCGATGA	60
	AACACTCATA ACTGAACACT GAGAGACTAT TCGCTTTCTC CTAAAGCCTT CAATCGAATT	120
15	CGCACG ATG AGA GGG AAC AAG TTT TGC TGT GAT TTC CGG TAC CTC CTC	168
	Met Arg Gly Asn Lys Phe Cys Cys Asp Phe Arg Tyr Leu Leu	
	450 455 460	
20	ATC TTG GCT GCT GTC GCC TTC ATC TAC ACA CAG ATG CGG CTT TTT GCG	216
	Ile Leu Ala Ala Val Ala Phe Ile Tyr Thr Gln Met Arg Leu Phe Ala	
	465 470 475	
25	ACA CAG TCA GAA TAT GCA GAT CGC CTT GCT GCT GCA ATT GAA GCA GAA	264
	Thr Gln Ser Glu Tyr Ala Asp Arg Leu Ala Ala Ala Ile Glu Ala Glu	
	480 485 490	
30	AAT CAT TGT ACA AGC CAG ACC AGA TTG CTT ATT GAC CAG ATT AGC CTG	312
	Asn His Cys Thr Ser Gln Thr Arg Leu Leu Ile Asp Gln Ile Ser Leu	
	495 500 505	
35	CAG CAA GGA AGA ATA GTT GCT CTT GAA GAA CAA ATG AAG CGT CAG GAC	360
	Gln Gln Gly Arg Ile Val Ala Leu Glu Glu Gln Met Lys Arg Gln Asp	
	510 515 520 525	
40	CAG GAG TGC CGA CAA TTA AGG GCT CTT GTT CAG GAT CTT GAA AGT AAG	408
	Gln Glu Cys Arg Gln Leu Arg Ala Leu Val Gln Asp Leu Glu Ser Lys	
	530 535 540	
45	GGC ATA AAA AAG TTG ATC GGA AAT GTA CAG ATG CCA GTG GCT GCT GTA	456
	Gly Ile Lys Lys Leu Ile Gly Asn Val Gln Met Pro Val Ala Ala Val	
	545 550 555	
50	GTT GTT ATG GCT TGC AAT CGG GCT GAT TAC CTG GAA AAG ACT ATT AAA	504
	Val Val Met Ala Cys Asn Arg Ala Asp Tyr Leu Glu Lys Thr Ile Lys	
	560 565 570	
55	TCC ATC TTA AAA TAC CAA ATA TCT GTT GCG TCA AAA TAT CCT CTT TTC	552
	Ser Ile Leu Lys Tyr Gln Ile Ser Val Ala Ser Lys Tyr Pro Leu Phe	
	575 580 585	
60	ATA TCC CAG GAT GGA TCA CAT CCT GAT GTC AGG AAG CTT GCT TTG AGC	600
	Ile Ser Gln Asp Gly Ser His Pro Asp Val Arg Lys Leu Ala Leu Ser	
	590 595 600 605	
65	TAT GAT CAG CTG ACG TAT ATG CAG CAC TTG GAT TTT GAA CCT GTG CAT	648
	Tyr Asp Gln Leu Thr Tyr Met Gln His Leu Asp Phe Glu Pro Val His	
	610 615 620	
70	ACT GAA AGA CCA GGG GAG CTG ATT GCA TAC TAC AAA ATT GCA CGT CAT	696
	Thr Glu Arg Pro Gly Glu Leu Ile Ala Tyr Tyr Lys Ile Ala Arg His	
	625 630 635	

# DE 197 54 622 A 1

TAC AAG TGG GCA TTG GAT CAG CTG TTT TAC AAG CAT AAT TTT AGC CGT	744	
Tyr Lys Trp Ala Leu Asp Gln Leu Phe Tyr Lys His Asn Phe Ser Arg		
640 645 650		
GTT ATC ATA CTA GAA GAT GAT ATG GAA ATT GCC CCT GAT TTT TTT GAC	792	5
Val Ile Ile Leu Glu Asp Asp Met Glu Ile Ala Pro Asp Phe Phe Asp		
655 660 665		
TTT TTT GAG GCT GGA GCT ACT CTT CTT GAC AGA GAC AAG TCG ATT ATG	840	10
Phe Phe Glu Ala Gly Ala Thr Leu Leu Asp Arg Asp Lys Ser Ile Met		
670 675 680 685		
GCT ATT TCT TCT TGG AAT GAC AAT GGA CAA ATG CAG TTT GTC CAA GAT	888	15
Ala Ile Ser Ser Trp Asn Asp Asn Gly Gln Met Gln Phe Val Gln Asp		
690 695 700		
CCT TAT GCT CTT TAC CGC TCA GAT TTT TTT CCC GGT CTT GGA TGG ATG	936	20
Pro Tyr Ala Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met		
705 710 715		
CTT TCA AAA TCT ACT TGG GAC GAA TTA TCT CCA AAG TGG CCA AAG GCT	984	25
Leu Ser Lys Ser Thr Trp Asp Glu Leu Ser Pro Lys Trp Pro Lys Ala		
720 725 730		
TAC TGG GAC GAC TGG CTA AGA CTC AAA GAG AAT CAC AGA GGT CGA CAA	1032	30
Tyr Trp Asp Asp Trp Leu Arg Leu Lys Glu Asn His Arg Gly Arg Gln		
735 740 745		
TTT ATT CGC CCA GAA GTT TGC AGA ACA TAT AAT TTT GGT GAG CAT GGT	1080	35
Phe Ile Arg Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly		
750 755 760 765		
TCT AGT TTG GGG CAG TTT TTC AAG CAG TAT CTT GAG CCA ATT AAA CTA	1128	40
Ser Ser Leu Gly Gln Phe Phe Lys Gln Tyr Leu Glu Pro Ile Lys Leu		
770 775 780		
AAT GAT GTC CAG GTT GAT TGG AAG TCA ATG GAC CTT AGT TAC CTT TTG	1176	45
Asn Asp Val Gln Val Asp Trp Lys Ser Met Asp Leu Ser Tyr Leu Leu		
785 790 795		
GAG GAC AAT TAC GTG AAA CAC TTT GGT GAC TTG GTT AAA AAG GCT AAG	1224	50
Glu Asp Asn Tyr Val Lys His Phe Gly Asp Leu Val Lys Lys Ala Lys		
800 805 810		
CCC ATC CAT GGA GCT GAT GCT GTC TTG AAA GCA TTT AAC ATA GAT GGT	1272	55
Pro Ile His Gly Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly		
815 820 825		
GAT GTG CGT ATT CAG TAC AGA GAT CAA CTA GAC TTT GAA AAT ATC GCA	1320	60
Asp Val Arg Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asn Ile Ala		
830 835 840 845		
CGG CAA TTT GGC ATT TTT GAA GAA TGG AAG GAT GGT GTA CCA CGT GCA	1368	65
Arg Gln Phe Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala		
850 855 860		
GCA TAT AAA GGA ATA GTA GTT TTC CGG TAC CAA ACG TCC AGA CGT GTA	1416	70
Ala Tyr Lys Gly Ile Val Val Phe Arg Tyr Gln Thr Ser Arg Arg Val		
865 870 875		

TTC CTT GTT GGC CAT GAT TCG CTT CAA CAA CTC GGA ATT GAA GAT ACT 1464  
 Phe Leu Val Gly His Asp Ser Leu Gln Gln Leu Gly Ile Glu Asp Thr  
           880                                  885                                  890

TAA CAAAGATATG ATTGCAGGAG CCCGGGCAAA ATTTTGGACT TATTGGGTAG 1517  
 \*

GATGCATCGA GCTGACACTA AACCATGATT TTACCAGTTA CATACAACGT TTTAATGTTA 1577

TACGGAGGAG CTCACTGTTC TAGTGTTGAA GGGATATCGG CTTCTTAGTA TTGGATGAAT 1637

CATCAACACA ACCTATTATT TTAAGTGTTT AGAACATAAA GAGGAAATGT AGCCCTGTAA 1697

AGACTATACA TGGGACCATC ATAATCGCGG CCGCGAATTC 1737

(2) ANGABEN ZU SEQ ID NO: 4:

(i) SEQUENZKENNZEICHEN:

(A) LÄNGE: 447 Aminosäuren

(B) ART: Aminosäure

(D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 4:

Met Arg Gly Asn Lys Phe Cys Cys Asp Phe Arg Tyr Leu Leu Ile Leu  
       1                                  5                                  10                                  15

Ala Ala Val Ala Phe Ile Tyr Thr Gln Met Arg Leu Phe Ala Thr Gln  
                                   20                                  25                                  30

Ser Glu Tyr Ala Asp Arg Leu Ala Ala Ala Ile Glu Ala Glu Asn His  
                                   35                                  40                                  45

Cys Thr Ser Gln Thr Arg Leu Ile Asp Gln Ile Ser Leu Gln Gln  
       50                                  55                                  60

Gly Arg Ile Val Ala Leu Glu Glu Gln Met Lys Arg Gln Asp Gln Glu  
       65                                  70                                  75                                  80

Cys Arg Gln Leu Arg Ala Leu Val Gln Asp Leu Glu Ser Lys Gly Ile  
                                   85                                  90                                  95

Lys Lys Leu Ile Gly Asn Val Gln Met Pro Val Ala Ala Val Val Val  
                                   100                                  105                                  110

Met Ala Cys Asn Arg Ala Asp Tyr Leu Glu Lys Thr Ile Lys Ser Ile  
                                   115                                  120                                  125

Leu Lys Tyr Gln Ile Ser Val Ala Ser Lys Tyr Pro Leu Phe Ile Ser  
                                   130                                  135                                  140

Gln Asp Gly Ser His Pro Asp Val Arg Lys Leu Ala Leu Ser Tyr Asp  
       145                                  150                                  155                                  160



## DE 197 54 622 A 1

Gln Leu Thr Tyr Met	Gln His Leu Asp	Phe Glu Pro Val	His Thr Glu	
165		170	175	
Arg Pro Gly Glu Leu Ile Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys				5
180	185		190	
Trp Ala Leu Asp Gln Leu Phe Tyr Lys His Asn Phe Ser Arg Val Ile				10
195	200		205	
Ile Leu Glu Asp Asp Met Glu Ile Ala Pro Asp Phe Phe Asp Phe Phe				
210	215		220	
Glu Ala Gly Ala Thr Leu Leu Asp Arg Asp Lys Ser Ile Met Ala Ile				15
225	230		235	240
Ser Ser Trp Asn Asp Asn Gly Gln Met Gln Phe Val Gln Asp Pro Tyr				20
245		250		255
Ala Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Ser				
260		265		270
Lys Ser Thr Trp Asp Glu Leu Ser Pro Lys Trp Pro Lys Ala Tyr Trp				25
275	280		285	
Asp Asp Trp Leu Arg Leu Lys Glu Asn His Arg Gly Arg Gln Phe Ile				30
290	295		300	
Arg Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly Ser Ser				
305	310		315	320
Leu Gly Gln Phe Phe Lys Gln Tyr Leu Glu Pro Ile Lys Leu Asn Asp				35
325		330		335
Val Gln Val Asp Trp Lys Ser Met Asp Leu Ser Tyr Leu Leu Glu Asp				40
340		345		350
Asn Tyr Val Lys His Phe Gly Asp Leu Val Lys Lys Ala Lys Pro Ile				
355	360		365	
His Gly Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly Asp Val				45
370	375		380	
Arg Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asn Ile Ala Arg Gln				50
385	390		395	400
Phe Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala Ala Tyr				
405		410		415
Lys Gly Ile Val Val Phe Arg Tyr Gln Thr Ser Arg Arg Val Phe Leu				55
420		425		430
Val Gly His Asp Ser Leu Gln Gln Leu Gly Ile Glu Asp Thr *				60
435	440		445	

(2) ANGABEN ZU SEQ ID NO: 5:

(i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 510 Basenpaare
- (B) ART: Nucleotid
- (C) STRANGFORM: Doppelstrang
- (D) TOPOLOGIE: linear

(ii) ART DES MOLEKÜLS: cDNA zu mRNA

(iii) HYPOTHETISCH: NEIN

(iv) ANTISENSE: NEIN

(vi) URSPRÜNGLICHE HERKUNFT:

- (A) ORGANISMUS: Arabidopsis thaliana
- (B) STAMM: Columbia
- (D) ENTWICKLUNGSSTADIUM: photosynthetisch aktive Organe
- (F) GEWEBETYP: Mesophyll
- (G) ZELLTYP: Blattzellen

(vii) UNMITTELBARE HERKUNFT:

- (B) CLON(E): pGEM-T/4a-3\* #12A

(ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: CDS
- (B) LÄGE: 1..510
- (D) SONSTIGE ANGABEN: /EC\_number= 2.4.1.101  
/product= "RT-PCR Produkt aus Blatt-mRNA"  
/gene= "cgl"  
/standard\_name= "gntI"  
/note= "homolog zu anderen GntI-Sequenzen"

(ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: primer\_bind
- (B) LÄGE: 1..48
- (D) SONSTIGE ANGABEN: /standard\_name= "PCR "sense"-Primer (gntI 4a) "

(ix) MERKMAL:

- (A) NAME/SCHLÜSSEL: primer\_bind
- (B) LÄGE: 487..510
- (D) SONSTIGE ANGABEN: /standard\_name= "PCR  
"antisense"-Primer (gntI 3\*) "

## (xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 5:

ATC GGA AAG CTT GGA TCC CCA GTG GCG GCT GTA GTT GTT ATG GCT TGC Ile Gly Lys Leu Gly Ser Pro Val Ala Ala Val Val Val Met Ala Cys 450 455 460	48	5
AGT CGT GCA GAC TAT CTT GAA AGG ACT GTT AAA TCA GTT TTA ACA TAT Ser Arg Ala Asp Tyr Leu Glu Arg Thr Val Lys Ser Val Leu Thr Tyr 465 470 475	96	10
CAA ACT CCC GTT GCT TCA AAA TAT CCT CTA TTT ATA TCT CAG GAT GGA Gln Thr Pro Val Ala Ser Lys Tyr Pro Leu Phe Ile Ser Gln Asp Gly 480 485 490 495	144	15
TCT GAT CAA GCT GTC AAG AGC AAG TCA TTG AGC TAT AAT CAA TTA ACA Ser Asp Gln Ala Val Lys Ser Lys Ser Leu Ser Tyr Asn Gln Leu Thr 500 505 510	192	20
TAT ATG CAG CAC TTG GAT TTT GAA CCA GTG GTC ACT GAA AGG CCT GGC Tyr Met Gln His Leu Asp Phe Glu Pro Val Val Thr Glu Arg Pro Gly 515 520 525	240	25
GAA CTG ACT GCG TAC TAC AAG ATT GCA CGT CAC TAC AAG TGG GCA CTG Glu Leu Thr Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys Trp Ala Leu 530 535 540	288	30
GAC CAG TTG TTT TAC AAA CAC AAA TTT AGT CGA GTG ATT ATA CTA GAA Asp Gln Leu Phe Tyr Lys His Lys Phe Ser Arg Val Ile Ile Leu Glu 545 550 555	336	35
GAT GAT ATG GAA ATT GCT CCA GAC TTC TTT GAT TAC TTT GAG GCT GCA Asp Asp Met Glu Ile Ala Pro Asp Phe Phe Asp Tyr Phe Glu Ala Ala 560 565 570 575	384	40
GCT AGT CTC ATG GAT AGG GAT AAA ACC ATT ATG GCT GCT TCA TCA TGG Ala Ser Leu Met Asp Arg Asp Lys Thr Ile Met Ala Ala Ser Ser Trp 580 585 590	432	45
ACT GAT AAT GGA CAG AAG CAG TTT GTG CAT GAT CCC TAT GCG CTA TAC Thr Asp Asn Gly Gln Lys Gln Phe Val His Asp Pro Tyr Ala Leu Tyr 595 600 605	480	50
CGA TCA GAT TTC TTC CCT GGC CAC GGC TGG Arg Ser Asp Phe Phe Pro Gly His Gly Trp 610 615	510	55

## (2) ANGABEN ZU SEQ ID NO: 6:

## (i) SEQUENZKENNZEICHEN:

- (A) LÄNGE: 170 Aminosäuren
- (B) ART: Aminosäure
- (D) TOPOLOGIE: linear

## (ii) ART DES MOLEKÜLS: Protein

(xi) SEQUENZBESCHREIBUNG: SEQ ID NO: 6:

5	Ile	Gly	Lys	Leu	Gly	Ser	Pro	Val	Ala	Ala	Val	Val	Val	Met	Ala	Cys
	1				5					10					15	
	Ser	Arg	Ala	Asp	Tyr	Leu	Glu	Arg	Thr	Val	Lys	Ser	Val	Leu	Thr	Tyr
				20					25					30		
10	Gln	Thr	Pro	Val	Ala	Ser	Lys	Tyr	Pro	Leu	Phe	Ile	Ser	Gln	Asp	Gly
			35					40					45			
	Ser	Asp	Gln	Ala	Val	Lys	Ser	Lys	Ser	Leu	Ser	Tyr	Asn	Gln	Leu	Thr
		50					55					60				
	Tyr	Met	Gln	His	Leu	Asp	Phe	Glu	Pro	Val	Val	Thr	Glu	Arg	Pro	Gly
	65					70					75					80
20	Glu	Leu	Thr	Ala	Tyr	Tyr	Lys	Ile	Ala	Arg	His	Tyr	Lys	Trp	Ala	Leu
				85					90						95	
	Asp	Gln	Leu	Phe	Tyr	Lys	His	Lys	Phe	Ser	Arg	Val	Ile	Ile	Leu	Glu
				100					105					110		
	Asp	Asp	Met	Glu	Ile	Ala	Pro	Asp	Phe	Phe	Asp	Tyr	Phe	Glu	Ala	Ala
			115					120					125			
30	Ala	Ser	Leu	Met	Asp	Arg	Asp	Lys	Thr	Ile	Met	Ala	Ala	Ser	Ser	Trp
			130				135					140				
	Thr	Asp	Asn	Gly	Gln	Lys	Gln	Phe	Val	His	Asp	Pro	Tyr	Ala	Leu	Tyr
	145					150					155					160
	Arg	Ser	Asp	Phe	Phe	Pro	Gly	His	Gly	Trp						
40				165					170							

## Patentansprüche

- 45 1. DNA-Sequenz, **dadurch gekennzeichnet**, daß sie eine pflanzliche N-Acetylglucosaminyltransferase I kodiert.
2. DNA-Sequenz, dadurch gekennzeichnet, daß sie N-Acetylglucosaminyltransferase I aus *Solanum tuberosum* kodiert.
3. DNA-Sequenz, dadurch gekennzeichnet, daß sie N-Acetylglucosaminyltransferase I aus *Nicotiana tabacum* kodiert.
- 50 4. DNA-Sequenz, dadurch gekennzeichnet, daß sie N-Acetylglucosaminyltransferase I aus *Arabidopsis thaliana* kodiert.
5. DNA-Sequenz nach Anspruch 4, dadurch gekennzeichnet, daß sie die in SEQ. ID. NO. 5 angegebene Nukleotidsequenz umfaßt.
6. DNA-Sequenz, dadurch gekennzeichnet, daß sie die in SEQ. ID. NO. 1 angegebene Nukleotidsequenz oder einen Teil davon oder eine entsprechende komplementäre Nukleotidsequenz umfaßt.
- 55 7. DNA-Sequenz, dadurch gekennzeichnet, daß sie die in SEQ. ID. NO. 3 angegebene Nukleotidsequenz oder einen Teil davon oder eine entsprechende komplementäre Nukleotidsequenz umfaßt.
8. DNA-Sequenz, dadurch gekennzeichnet, daß sie die in SEQ. ID. NO. 5 angegebene Nukleotidsequenz oder einen Teil davon oder eine entsprechende komplementäre Nukleotidsequenz umfaßt.
- 60 9. DNA-Sequenz, dadurch gekennzeichnet, daß sie durch Substitution, Deletion und/oder Insertion einzelner oder mehrerer Nukleotide und/oder Verkürzung und/oder Verlängerung am 5'- und/oder 3'-Ende einer der DNA-Sequenzen nach einem der Ansprüche 1 bis 8 erhalten wird mit der Maßgabe, daß die DNA-Sequenz unter stringenten Bedingungen zumindest in einem Teilabschnitt mit der Ausgangs-DNA-Sequenz oder deren komplementärer Sequenz oder mit Teilen derselben hybridisiert.
- 65 10. DNA-Sequenz, dadurch gekennzeichnet, daß sie ein Gen darstellt oder Teil eines Gens ist, das das Enzym N-Acetylglucosaminyltransferase I kodiert, und in ihrer Gesamtheit oder in einem Teilabschnitt mit einer der DNA-Sequenzen oder -Fragmente nach einem der Ansprüche 1 bis 9 und/oder mit einer DNA-Sequenz, die von den in den SEQ. ID. NO.: 1, 3 und/oder 5 angegebenen Aminosäuresequenzen

zen unter Berücksichtigung der Degeneration des genetischen Codes abgeleitet worden ist, unter stringenten Bedingungen hybridisiert.

11. DNA-Sequenz, dadurch gekennzeichnet, daß sie in ihrer Gesamtheit oder in einem Teilbereich mit einer oder mehreren der DNA-Sequenzen nach einem der Ansprüche 1 bis 10 hybridisiert.
12. DNA-Konstrukt, dadurch gekennzeichnet, daß es eine oder mehrere der DNA-Sequenzen gemäß einem der Ansprüche 1 bis 11 umfaßt.
13. "antisense"-Konstrukt, dadurch gekennzeichnet, daß es eine "antisense"-DNA bezüglich der DNA-Sequenz nach einem der Ansprüche 1 bis 12 und ggf. regulatorische Sequenzen für die Transkription der "antisense"-DNA umfaßt.
14. "sense"-Konstrukt, dadurch gekennzeichnet, daß es eine "sense"-DNA bezüglich der DNA-Sequenz nach einem der Ansprüche 1 bis 12 und ggf. regulatorische Sequenzen für die Transkription der "sense"-DNA umfaßt.
15. Vektor, Plasmid, Cosmid, Viren- oder Phagengenom, dadurch gekennzeichnet, daß er oder es zumindest eine DNA-Sequenz und/oder ein Konstrukt nach einem der Ansprüche 1 bis 14 umfaßt.
16. Pflanzliche N-Acetylglucosaminyltransferase I
17. N-Acetylglucosaminyltransferase I aus *Solanum tuberosum*.
18. N-Acetylglucosaminyltransferase I aus *Nicotiana tabacum*.
19. N-Acetylglucosaminyltransferase I aus *Arabidopsis thaliana*.
20. N-Acetylglucosaminyltransferase I nach Anspruch 19, dadurch gekennzeichnet, daß das Enzym die in SEQ ID NO: 6 angegebene Aminosäuresequenz umfaßt.
21. N-Acetylglucosaminyltransferase I, dadurch gekennzeichnet, daß das Enzym die in SEQ ID NO: 2 angegebene Aminosäuresequenz umfaßt.
22. N-Acetylglucosaminyltransferase I, dadurch gekennzeichnet, daß das Enzym die Aminosäuren 74 bis 446 der SEQ ID NO: 2 angegebenen Aminosäuresequenz umfaßt.
23. N-Acetylglucosaminyltransferase I, dadurch gekennzeichnet, daß das Enzym die in SEQ ID NO: 4 angegebene Aminosäuresequenz umfaßt.
24. N-Acetylglucosaminyltransferase I, zugänglich aufgrund der Hybridisierung ihres Gen oder eines oder mehrerer Abschnitte ihres Gens mit einer oder mehreren der DNA-Sequenzen und/oder DNA-Fragmente gemäß einem der Ansprüche 1 bis 11.
25. Von den Enzymen gemäß einem der Ansprüche 16 bis 24 durch Substitution, Deletion, Insertion und/oder Modifikation einzelner Aminosäuren und/oder kleinerer Gruppen von Aminosäuren und/oder durch N- und/oder C-terminale Verkürzung und/oder Verlängerung abgeleitete Enzyme oder Proteine.
26. Protein oder Peptid, umfassend einen oder mehrere Abschnitte der Aminosäuresequenz(en) eines oder mehrerer der in einem der Ansprüche 16 bis 25 definierten Enzyme.
27. Protein oder Peptid, welches durch eine der DNA-Sequenzen nach einem der Ansprüche 1 bis 11 kodiert wird.
28. Antigen, dadurch gekennzeichnet, daß es
  - die in SEQ ID NO: 2, SEQ ID NO: 4 oder SEQ ID NO: 6 angegebene Aminosäuresequenz oder
  - die Aminosäuren 74 bis 446 der in Fig. 2 angegebenen Aminosäuresequenz oder
  - eine durch Substitution, Deletion, Insertion und/oder Modifikation einzelner Aminosäuren und/oder kleinerer Gruppen von Aminosäuren von den in SEQ ID NO: 2, 4 oder 6 angegebenen Aminosäuresequenzen abgeleitete Aminosäuresequenz oder
  - einen Teil oder mehrere Teile dieser Sequenzen
 umfaßt mit der Maßgabe, daß das Antigen bei Immunisierung eines Wirtes mit diesem zur Bildung einer immunologischen Reaktion einschließlich der Erzeugung von gegen das Antigen gerichteten Antikörpern führt.
29. Monoklonaler oder polyklonaler Antikörper, dadurch gekennzeichnet, daß er spezifisch eines oder mehrere der Enzyme oder Antigene nach einem der Ansprüche 16 bis 28 erkennt und dieses oder diese bindet.
30. Monoklonaler oder polyklonaler Antikörper, dadurch gekennzeichnet, daß er durch Immunisierung eines Wirtes mit einem oder mehreren Antigenen nach Anspruch 28 erhalten werden kann.
31. Mikroorganismus, dadurch gekennzeichnet, daß er durch mindestens eine der Nukleotidsequenzen, ausgewählt aus den DNA-Sequenzen, Konstrukten, Vektoren, Plasmiden, Cosmiden, Viren- oder Phagengenomen gemäß einem oder mehreren der Ansprüche 1 bis 15, transformiert ist.
32. Verwendung eines oder mehrerer Antikörper nach Anspruch 29 und/oder Anspruch 30 zur Detektion von Pflanzen mit fehlendem oder defektem N-Acetylglucosaminyltransferase I-Enzym.
33. Verwendung einer oder mehrerer der DNA-Sequenzen und/oder Konstrukte nach einem der Ansprüche 1 bis 15 zur Detektion von Pflanzen mit fehlendem oder defektem Gen für das N-Acetylglucosaminyltransferase I-Enzym.
34. Verwendung eines oder mehrerer "antisense"-Konstrukte nach Anspruch 13 zur Erzeugung von Pflanzen mit fehlender oder verminderter N-Acetylglucosaminyltransferase I-Aktivität.
35. Verwendung eines oder mehrerer "sense"-Konstrukte nach Anspruch 14 zur Erzeugung von Pflanzen mit fehlender oder verminderter N-Acetylglucosaminyltransferase I-Aktivität.
36. Transgene Pflanze, transgener Samen, transgenes Vermehrungsmaterial, Teile transgener Pflanzen oder transformierte Pflanzenzelle, erhältlich durch Integration einer oder mehrerer DNA-Sequenz(en) oder Konstrukt(e) nach einem der Ansprüche 1 bis 12 unter Kontrolle eines in Pflanzen wirksamen Promotors in das Genom einer Pflanze oder durch Infektion durch ein eine oder mehrere DNA-Sequenz(en) oder Konstrukt(e) nach einem der Ansprüche 1 bis 12 enthaltendes Virus für eine extrachromosomale Propagation und Expression der DNA-Sequenz(en) oder des Konstrukts/der Konstrukte in infiziertem Pflanzengewebe.
37. Transgene Pflanze, transgener Samen, transgenes Vermehrungsmaterial, Teile transgener Pflanzen oder transformierte Pflanzenzelle mit fehlender oder verminderter N-Acetylglucosaminyltransferase I-Aktivität, erhältlich durch Integration eines oder mehrerer "antisense"-Konstrukte nach Anspruch 13 unter Kontrolle eines in Pflanzen wirksamen Promotors in das Genom einer Pflanze oder durch virale Infektion durch ein eines oder mehrere "anti-

sense"-Konstrukte) nach Anspruch 13 enthaltendes Virus für eine extrachromosomale Propagation und Transkription des "antisense"-Konstrukts oder der "antisense"-Konstrukte in infiziertem Pflanzengewebe.

38. Transgene Pflanze, transgener Samen, transgenes Vermehrungsmaterial, Teile transgener Pflanzen oder transformierte Pflanzenzelle mit fehlender oder verminderter N-Acetylglucosaminyltransferase I-Aktivität, erhältlich durch Integration eines oder mehrerer "sense"-Konstrukte) nach Anspruch 14 unter Kontrolle eines in Pflanzen wirksamen Promotors in das Genom einer Pflanze oder durch virale Infektion durch ein eines oder mehrere "sense"-Konstrukte) nach Anspruch 14 enthaltendes Virus für eine extrachromosomale Propagation und Transkription des "sense"-Konstrukts oder der "sense"-Konstrukte in infiziertem Pflanzengewebe.

39. Verfahren zur Erzeugung von Glykoproteinen mit minimalen, einheitlichen und definierten Zuckerresten, umfassend das Züchten einer transgenen Pflanze, von Teilen transgener Pflanzen oder von transformierten Pflanzenzellen nach Anspruch 37 oder 38 und das Isolieren des gesuchten Glykoproteins aus dem gezüchteten Material.

40. Verfahren nach Anspruch 39, dadurch gekennzeichnet, daß die eingesetzte transgene Pflanze zusätzlich mit dem das gesuchte Glykoprotein kodierenden Gen transformiert worden ist.

41. Verfahren nach Anspruch 40, dadurch gekennzeichnet, daß die Transformation der eingesetzten transgenen Pflanze durch Integration des das gesuchte Glykoprotein kodierenden Gens unter Kontrolle eines in Pflanzen wirksamen Promotors in das Genom der transgenen Pflanze oder durch eine Infektion der transgenen Pflanze durch ein Virus, das ein Gen enthält, das das gesuchte Glykoprotein kodiert, für eine extrachromosomale Propagation und Expression des Gens in infiziertem Pflanzengewebe erfolgt.

42. Verfahren nach Anspruch 40, dadurch gekennzeichnet, daß die zusätzlich mit dem das gesuchte Glykoprotein kodierenden Gen transformierte transgene Pflanze durch gleichzeitige Transformation einer Ausgangspflanze sowohl mit dem "antisense"- oder "sense"-Konstrukt als auch dem das gesuchte Glykoprotein kodierenden Gen erzeugt wurde.

43. Verfahren nach Anspruch 42, dadurch gekennzeichnet, daß die Transformation durch virale Infektion durch ein Virus, das sowohl das "antisense"- oder "sense"-Konstrukt als auch das das gesuchte Glykoprotein kodierende Gen enthält, für eine extrachromosomale Propagation und Transkription oder Expression in infiziertem Pflanzengewebe erfolgt.

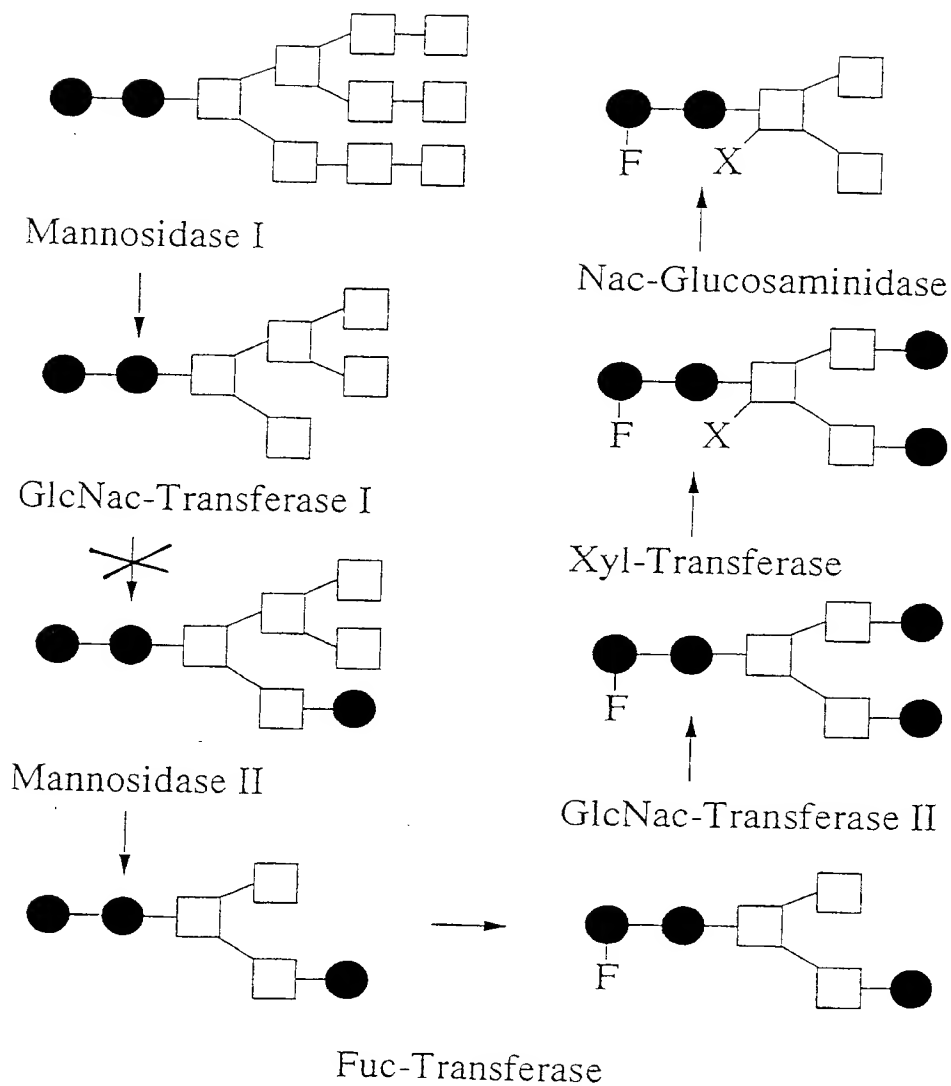
---

Hierzu 5 Seiten(n) Zeichnungen

---

- Leerseite -

Figur 1





Figur 2

A1 *GntI* cDNA

```

GAATTCGCGG CCGCCTGAGA AACCCCTCGAA TTCAATTTCG CATTGGGCAG AG ATG      55
                                     Met
                                     1

AGA GGG AAC AAG TTT TGC TTT GAT TTA CGG TAC CTT CTC GTC GTG GCT      103
Arg Gly Asn Lys Phe Cys Phe Asp Leu Arg Tyr Leu Leu Val Val Ala
                    5                      10                      15

GCT CTC GCC TTC ATC TAC ATA CAG ATG CGG CTT TTC GCG ACA CAG TCA      151
Ala Leu Ala Phe Ile Tyr Ile Gln Met Arg Leu Phe Ala Thr Gln Ser
                20                      25                      30

GAA TAT GTA GAC CGC CTT GCT GCT GCA ATT GAA GCA GAA AAT CAT TGT      199
Glu Tyr Val Asp Arg Leu Ala Ala Ile Glu Ala Glu Asn His Cys
                35                      40                      45

ACA AGT CAG ACC AGA TTG CTT ATT GAC AAG ATT AGC CAG CAG CAA GGA      247
Thr Ser Gln Thr Arg Leu Leu Ile Asp Lys Ile Ser Gln Gln Gln Gly
                50                      55                      60                      65

AGA GTA GTA GCT CTT GAA GAA CAA ATG AAG CAT CAG GAC CAG GAG TGC      295
Arg Val Val Ala Leu Glu Glu Gln Met Lys His Gln Asp Gln Glu Cys
                70                      75                      80

CGG CAA TTA AGG GCT CTT GTT CAG GAT CTT GAA AGT AAG GGC ATA AAA      343
Arg Gln Leu Arg Ala Leu Val Gln Asp Leu Glu Ser Lys Gly Ile Lys
                85                      90                      95

AAG TTA ATC GGA GAT GTG CAG ATG CCA GTG GCA GCT GTA GTT GTT ATG      391
Lys Leu Ile Gly Asp Val Gln Met Pro Val Ala Ala Val Val Val Met
                100                      105                      110

GCT TGC AGT CGT ACT GAC TAC CTG GAG AGG ACT ATT AAA TCC ATC TTA      439
Ala Cys Ser Arg Thr Asp Tyr Leu Glu Arg Thr Ile Lys Ser Ile Leu
                115                      120                      125

AAA TAC CAA ACA TCT GTT GCA TCA AAA TAT CCT CTT TTC ATA TCC CAG      487
Lys Tyr Gln Thr Ser Val Ala Ser Lys Tyr Pro Leu Phe Ile Ser Gln
                130                      135                      140                      145

GAT GGA TCA AAT CCT GAT GTA AGA AAG CTT GCT TTG AGC TAT GGT CAG      535
Asp Gly Ser Asn Pro Asp Val Arg Lys Leu Ala Leu Ser Tyr Gly Gln
                150                      155                      160

CTG ACG TAT ATG CAG CAC TTG GAT TAT GAA CCT GTG CAT ACT GAA AGA      583
Leu Thr Tyr Met Gln His Leu Asp Tyr Glu Pro Val His Thr Glu Arg
                165                      170                      175

CCA GGG GAA CTG GTT GCA TAC TAC AAG ATT GCA CGT CAT TAC AAG TGG      631
Pro Gly Glu Leu Val Ala Tyr Tyr Lys Ile Ala Arg His Tyr Lys Trp
                180                      185                      190

GCA TTG GAT CAG CTG TTT CAC AAG CAT AAT TTT AGC CGT GTT ATC ATA      679
Ala Leu Asp Gln Leu Phe His Lys His Asn Phe Ser Arg Val Ile Ile
                195                      200                      205

CTA GAA GAT GAT ATG GAA ATT GCT GCT GAT TTT TTT GAC TAT TTT GAG      727
Leu Glu Asp Asp Met Glu Ile Ala Ala Asp Phe Phe Asp Tyr Phe Glu
                210                      215                      220                      225

```

GCT GGA GCT ACT CTT CTT GAC AGA GAC AAG TCG ATT ATG GCT ATT TCT Ala Gly Ala Thr Leu Leu Asp Arg Asp Lys Ser Ile Met Ala Ile Ser 230 235 240	775
TCT TGG AAT GAC AAT GGA CAA AGG CAG TTC GTC CAA GAT CCT GAT GCT Ser Trp Asn Asp Asn Gly Gln Arg Gln Phe Val Gln Asp Pro Asp Ala 245 250 255	823
CTT TAC CGC TCA GAC TTT TTT CCT GGT CTT GGA TGG ATG CTT TCA AAA Leu Tyr Arg Ser Asp Phe Phe Pro Gly Leu Gly Trp Met Leu Ser Lys 260 265 270	871
TCA ACT TGG TCC GAA CTA TCT CCA AAG TGG CCA AAG GCT TAC TGG GAT Ser Thr Trp Ser Glu Leu Ser Pro Lys Trp Pro Lys Ala Tyr Trp Asp 275 280 285	919
GAC TGG CTA AGG CTG AAA GAA AAT CAC AGA GGT CGA CAA TTT ATT CGC Asp Trp Leu Arg Leu Lys Glu Asn His Arg Gly Arg Gln Phe Ile Arg 290 295 300 305	967
CCA GAA GTT TGC AGA ACG TAC AAT TTT GGT GAG CAT GGT TCT AGT TTG Pro Glu Val Cys Arg Thr Tyr Asn Phe Gly Glu His Gly Ser Ser Leu 310 315 320	1015
GGG CAG TTT TTT AAG CAG TAT CTT GAG CCA ATT AAG CTA AAT GAT GTC Gly Gln Phe Phe Lys Gln Tyr Leu Glu Pro Ile Lys Leu Asn Asp Val 325 330 335	1063
CAG GTT GAT TGG AAG TCA ATG GAC CTA AGT TAC CTT TTG GAG GAC AAC Gln Val Asp Trp Lys Ser Met Asp Leu Ser Tyr Leu Leu Glu Asp Asn 340 345 350	1111
TAT GTG AAA CAC TTT GGC GAC TTG GTT AAA AAG GCT AAG CCC ATC CAC Tyr Val Lys His Phe Gly Asp Leu Val Lys Lys Ala Lys Pro Ile His 355 360 365	1159
GGA GCT GAT GCT GTT TTG AAA GCA TTT AAC ATA GAT GGT GAT GTG CGT Gly Ala Asp Ala Val Leu Lys Ala Phe Asn Ile Asp Gly Asp Val Arg 370 375 380 385	1207
ATT CAG TAC AGA GAC CAA CTA GAC TTT GAA GAT ATC GCT CGA CAG TTT Ile Gln Tyr Arg Asp Gln Leu Asp Phe Glu Asp Ile Ala Arg Gln Phe 390 395 400	1255
GGC ATT TTT GAA GAA TGG AAG GAT GGT GTA CCA CGG GCA GCA TAT AAA Gly Ile Phe Glu Glu Trp Lys Asp Gly Val Pro Arg Ala Ala Tyr Lys 405 410 415	1303
GGG ATA GTA GTT TTC CGG TTT CAA ACA TCT AGA CGT GTG TTC CTT GTT Gly Ile Val Val Phe Arg Phe Gln Thr Ser Arg Arg Val Phe Leu Val 420 425 430	1351
TCC CCT GAT TCT CTT CGA CAA CTT GGA GTT GAA GAT ACT TAG Ser Pro Asp Ser Leu Arg Gln Leu Gly Val Glu Asp Thr End 435 440 445	1393
CGAAGATATG ATTGGAGCCT GAGCAACAAT TTAGACTTAT TTGGTAGGAT ACATTTGAAA	1453
GAGCTGACAC GAAAAGTATG ACTACCAGTA GCTACATGCA ACATTTTAAT GTTAATGGAA	1513
GGAACCCACT GCTTATTGTT GGAATGGATG AATCATCACC ACATCCTATT ATTCAAGTTT	1573
ACAAACATAA AGAGGAAATG TTGCCCTATA AAAACAAATT TTTTGTTTCT AAGAAGGAAC	1633
GTTACGATTA TGAGCAACTT TGGCGGCCGC GAATTC	1669

Figur 3A

Hu	Ra	Mo	Ce	
35 (59)	36 (57)	35 (59)	33 (57)	St
	92 (95)	91 (94)	38 (57)	Hu
		90 (93)	38 (57)	Ra
			38 (58)	Mo

Figur 3B

A\_Stb-A1 1 MRGNKFCFDLRYLLVVAALAFIYIOMRLFATQSEYVDRLAAAIEAENHCT  
 B\_Ntb-A9 1 MRGNKFCDFRYLLILAAVAFIYIOMRLFATQSEYADR LAAAIEAENHCT  
 C\_Atb-Pcr 1 .....

A\_Stb-A1 51 SQTRLIDKISLQOQGRVVALEEQMKHODQECROLRALVQDLESKGIKKLI  
 B\_Ntb-A9 51 SQTRLIDQISLQOQGRIVALEEQMKRQDQECROLRALVQDLESKGIKKLI  
 C\_Atb-Pcr 1 .....

A\_Stb-A1 101 GNVQMPVAAVVVMACSEEDYLERTIKSILKYQTSVASKYPLFISQDGSNP  
 B\_Ntb-A9 101 GNVQMPVAAVVVMACNRADYLEKTIKSILKYQTSVASKYPLFISQDGSHP  
 C\_Atb-Pcr 1 .....SRADYLERTVKSVLTYQTPVASKYPLFISQDGSQDQ

A\_Stb-A1 151 DVRKLALSYGQLTVMQHLDFEPVTERPGELVAYYKIARHYKWALDQLFH  
 B\_Ntb-A9 151 DVRKLALSVDQLTVMQHLDFEPVTERPGELVAYYKIARHYKWALDQLFY  
 C\_Atb-Pcr 36 AVKSKSLSYNQLTVMQHLDFEPVTERPGELVAYYKIARHYKWALDQLFY

A\_Stb-A1 201 KHNF SRVILEDDMEIADFFDYFEAGATLLDRDKSIMAISWNDNGORQ  
 B\_Ntb-A9 201 KHNF SRVILEDDMEIADFFDYFEAGATLLDRDKSIMAISWNDNGOMQ  
 C\_Atb-Pcr 86 KHNF SRVILEDDMEIADFFDYFEAAASLMDRDKTIMAASSWTDNGOKO

A\_Stb-A1 251 FVQDPDALYRSDFFPGLGWMLSKSTWS ELSPKWPKAYWDDWLRLKENHRG  
 B\_Ntb-A9 251 FVQDPYALYRSDFFPGLGWMLSKSTWDELSPKWPKAYWDDWLRLKENHRG  
 C\_Atb-Pcr 136 FVHDPYALYRS .....

A\_Stb-A1 301 RQFIRPEVCRTYNFGEHGSSLGQFFKOYLEPIKLNDVQVDWKSMDLSYLL  
 B\_Ntb-A9 301 RQFIRPEVCRTYNFGEHGSSLGQFFKOYLEPIKLNDVQVDWKSMDLSYLL  
 C\_Atb-Pcr .....

A\_Stb-A1 351 EDNYVKHFGDLVKKAKPIHGADAVLKAENIDGDVRIQYRDQLDFFEDTARQ  
 B\_Ntb-A9 351 EDNYVKHFGDLVKKAKPIHGADAVLKAENIDGDVRIQYRDQLDFFENIARQ  
 C\_Atb-Pcr .....

A\_Stb-A1 401 FGIFEWKDGVPRAAAYKGIVVFRYOTSRRVFLVSPDSLROLGVEDT  
 B\_Ntb-A9 401 FGIFEWKDGVPRAAAYKGIVVFRYOTSRRVFLVGHDSLQOLGIEDT  
 C\_Atb-Pcr .....

Figur 4

